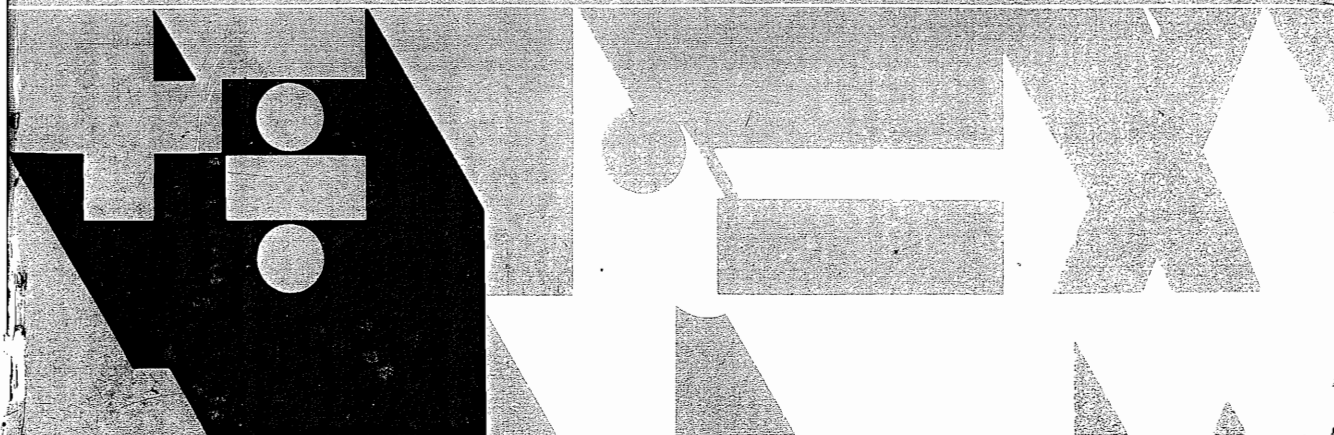


MATHEMATICAL ANALYSIS

BUSINESS AND ECONOMIC
APPLICATIONS

THIRD EDITION

HARPER & ROW, PUBLISHERS, INC.

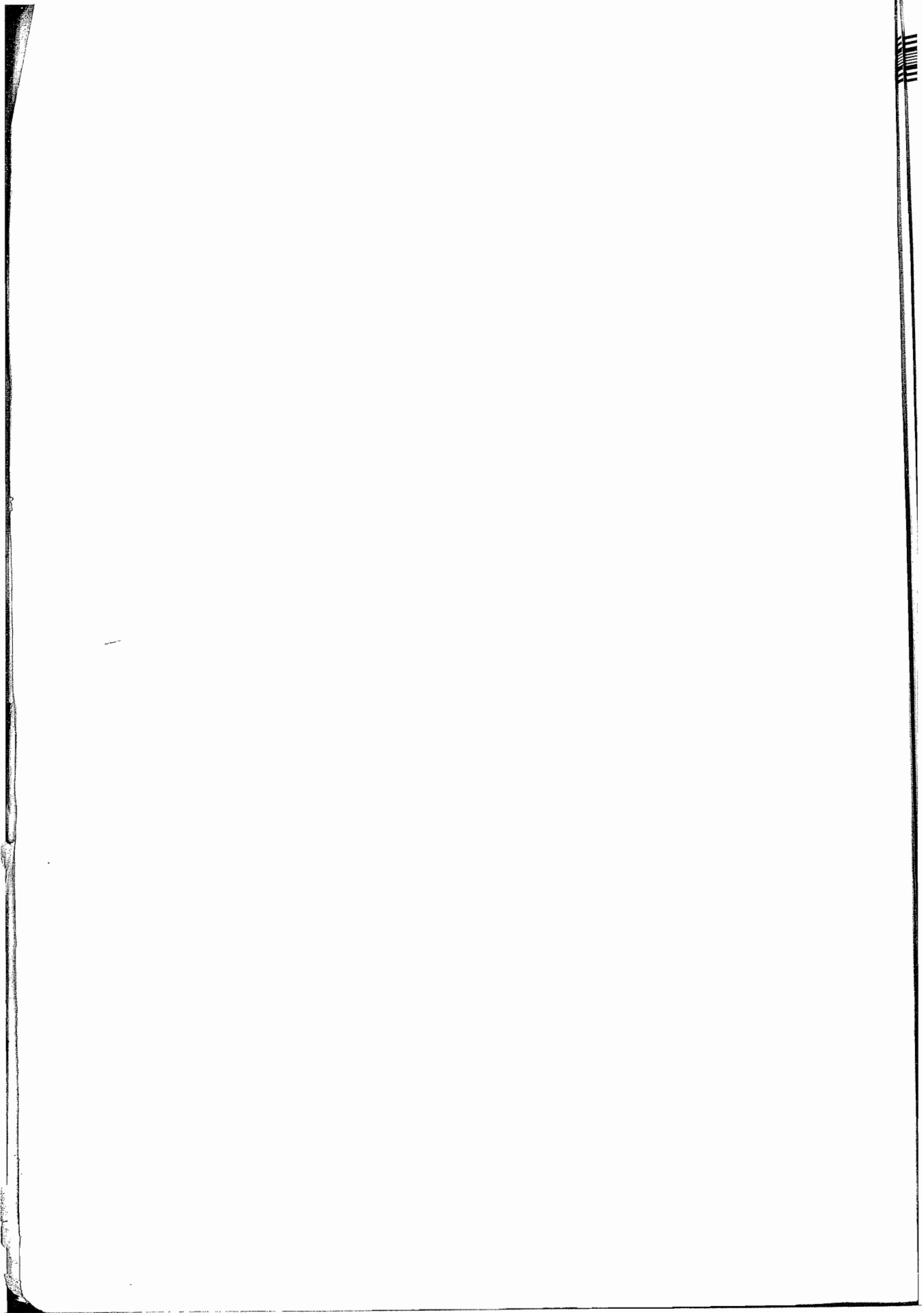


**MATHEMATICAL
ANALYSIS**

**EX LIBRIS PROF. DR.
DARCY CARVALHO**



SINE ME LIBER IBIS



Contents

Preface to the Third Edition, xi
Preface, xii

INTRODUCTION, 1

Sets, 2
Variables, 6
Relations and Functions, 8
Inverse Functions, 13

1. GRAPHICAL REPRESENTATION, 16

- 1.1 Introduction, 16
- 1.2 Rectangular Coordinates, 16
 - Fundamental Principle, 17
- 1.3 Straight Lines, 17
 - Slope of a Straight Line, 19
 - Equations for Straight Lines, 21
 - Two-Point Form, 21 Intercept Form, 22 Point-Slope Form, 23
 - Slope-Intercept Form, 24 Vertical Lines, 24 Horizontal Lines, 25
 - Parallel, Perpendicular, and Intersecting Lines, 26
 - Intersection of Two Straight Lines, 27
 - Independent Equations, 27 Consistent Equations, 27
 - Families of Lines, 33
- 1.4 Applications of Straight Lines in Business and Economics, 37
 - Linear Demand and Supply Curves, 37
 - Linear Demand Curves, 38 Linear Supply Curves, 40
 - Market Equilibrium, 43
 - Break-Even Analysis, 44
 - The Consumption Function, 46
- 1.5 General Methods of Graphing Nonlinear Curves, 50
 - Intercepts, 51
 - Symmetry, 51
 - Extent, 56
 - Asymptotes, 57
 - Factorization, 60
 - Real Curve, Point Locus, or Imaginary Locus, 61
 - Summary: General Methods of Graphing Nonlinear Curves, 62
- 1.6 Quadratic Curves, 70
 - Identification of a Quadratic Equation, 72

- The Circle, 72
 - The Ellipse, 74
 - The Parabola, 77
 - The Hyperbola, 81
 - Special Cases of the Equilateral Hyperbola, 86
 - Families of Quadratic Curves, 90
 - 1.7 Applications of Quadratic Curves in Business and Economics, 92**
 - Demand and Supply Curves, 92
 - Market Equilibrium, 93
 - Product Transformation Curves, 98
 - Pareto's Law of Distribution of Income, 107
 - 1.8 Exponential and Logarithmic Curves, 111**
 - Exponential Functions, 111
 - Logarithmic Functions, 114
 - 1.9 Applications of Exponential and Logarithmic Curves in Business and Economics, 117**
 - Compound Interest, 118
 - Growth Functions, 119
 - Biological Growth Curves, 119
 - Gompertz Curves, 120
 - Learning Curves, 121
 - 1.10 Trigonometric Curves, 123**
 - Degree and Radian Measure, 124
 - Trigonometric Functions, 125
 - Trigonometric Identities, 129
 - Polar Coordinates, 129
 - 1.11 Applications of Trigonometric Curves in Business and Economics, 133**
 - Technical Notes I, II, III, IV, and V 133, 133, 134, 137, and 137*
- 2. DIFFERENTIAL CALCULUS: FUNCTIONS OF ONE VARIABLE, 139**
- 2.1 Introduction, 139**
 - 2.2 Limits, 141**
 - Definition of a Limit, 141
 - Properties of Limits, 146
 - 2.3 Continuity, 151**
 - Types of Discontinuities, 151
 - Properties of Continuous Functions, 159
 - 2.4 Definition of the First Derivative, 161**
 - Interpretation of the First Derivative, 165
 - Velocity of a Moving Body, 165
 - Rate of Change of a Function, 167
 - 2.5 Applications of the First Derivative in Business and Economics, 168**
 - Cost, 168
 - Revenue, 170
 - Elasticity, 171
 - National Income, Consumption, and Savings, 172
 - 2.6 Rules for Differentiation, 174**
 - Algebraic Functions, 174
 - Rule for Differentiation of Polynomial Functions, 177
 - Logarithmic Functions, 185
 - Exponential Functions, 187
 - Trigonometric Functions, 191
 - Inverse Functions, 194
 - Composite Functions, 196
 - Summary of Procedure for Differentiation, 199
 - Summary of Rules for Differentiation, 201
 - Algebraic Functions, 201
 - Logarithmic Functions, 202
 - Exponential Functions, 202
 - Trigonometric Functions, 202
 - Inverse Functions, 202
 - Composite Functions, 202
 - 2.7 Differentials, 203**
 - Approximations, 204
 - 2.8 Higher-Order Derivatives, 207**

- 2.9 Implicit Differentiation, 213**
 Procedure for Implicit Differentiation, 213
 - 2.10 Differentiability and Continuity, 217**
 Smooth Functions, 218
 - 2.11 Applications of Derivatives, 219**
 The First Derivative, 219
 Increasing and Decreasing Functions, 219 Relative (Local) Maxima and Minima, 220
 Summary of the Information That Can Be Obtained from the First Derivative, 230
 The Second Derivative, 231
 Concavity, 231 Points of Inflection, 236
 Summary of Information That Can Be Obtained from the Second Derivative, 240
 Summary of the Procedure for Sketching the Graph of $y = f(x)$
 Using Information Provided by First and Second Derivatives, 240
 - 2.12 Applications of Derivatives in Business and Economics, 245**
 Cost, Average Cost, and Marginal Cost, 245
 Elasticity, 256
 Constant Elasticity of Demand, 263
 Revenue, Marginal Revenue, and Elasticity of Demand, 266
 Revenue from Taxation, 271
 Profit Under Monopoly, 275
 Effect of Taxation on Monopoly, 280
 Inventory Models, 285
 - 2.13 Indeterminate Forms, 289**
 - 2.14 Sequences and Series, 298**
 Summary of Procedure for Testing Convergence or Divergence of an Infinite Series, 300
 Power Series, 304
 Taylor's Theorem, 308
 Technical Notes I, II, and III 317, 321, and 321
- 3. DIFFERENTIAL CALCULUS: FUNCTIONS OF MORE THAN ONE VARIABLE, 323**
- 3.1 Introduction, 323**
 Functions of More Than One Variable, 323
 Continuity, 324
 - 3.2 Partial Differentiation, 324**
 The Total Differential, 330
 The Total Derivative, 332
 Differentiation of Implicit Functions, 335
 - 3.3 Applications of Partial Derivatives in Business and Economics, 340**
 Marginal Cost, 340
 Demand Surfaces, 341
 Marginal Demand, 341 Partial Elasticities of Demand, 344
 Production Functions, 346
 Marginal Productivity, 346
 Euler's Theorem, 348
 Linear Homogeneous Production Functions, 349
 Constant Product Curves, 349
 Returns to Scale, 351
 Utility Functions, 352
 - 3.4 Maxima and Minima of Functions of Two Variables, 356**
 - 3.5 Maxima and Minima Subject to Constraints, 367**
 Lagrange Multipliers, 368
 Kuhn-Tucker Conditions, 375
 - 3.6 Applications of Constrained Maximization and Minimization in Business and Economics, 380**
 Constrained Utility Maximization, 380
 Technical Note I 385

4. INTEGRAL CALCULUS, 387

- 4.1 Introduction, 387**
- 4.2 Indefinite Integration, 388**
- 4.3 Applications of Indefinite Integration in Business and Economics, 394**
 - Cost, 394
 - Revenue, 395
 - National Income, Consumption, and Savings, 396
 - Capital Formation, 397
- 4.4 Definite Integration, 399**
 - Area as a Definite Integral, 404
 - Interpretation of Negative Areas, 407
 - Area Between Two Curves, 408
 - Probability as an Area, 414
- 4.5 Applications of Definite Integration in Business and Economics, 418**
 - Consumers' Surplus, 418
 - Producers' Surplus, 422
 - Revenue Versus Cost, 425
- 4.6 Special Methods of Integration, 431**
 - Standard Forms for Integration, 431
 - Integration by Parts, 435
 - Integration by Partial Fractions, 438
 - Integration by a Rationalizing Substitution, 443
 - Integration by Miscellaneous Substitutions, 447
- 4.7 Numerical (Approximate) Methods of Integration, 449**
 - The Trapezoidal Rule, 449
 - Notes Concerning the Accuracy of the Trapezoidal Rule, 450
 - Simpson's Rule, 451
 - Taylor's Expansion, 453
- 4.8 Multiple Integrals, 459**
 - Probability as a Double Integral, 464

5. DIFFERENTIAL EQUATIONS, 469

- 5.1 Introduction, 469**
- 5.2 Definition and Classification of Differential Equations, 470**
- 5.3 Solutions of Ordinary Differential Equations, 471**
- 5.4 Differential Equations of First Order and First Degree, 476**
 - Separable Differential Equations, 477
 - Homogeneous Differential Equations, 480
 - Exact Differential Equations, 487
 - Linear Differential Equations, 491
 - Differential Equations Linear in a Function of y or in a Function of x , 495
- 5.5 Differential Equations of Higher Order and/or Higher Degree, 498**
- 5.6 Applications of Differential Equations in Economic Models, 502**
 - Domar Macro Model, 502
 - Domar Debt Model, 503
 - A Second Domar Debt Model, 504
 - Evans Price Adjustment Model, 505
 - Income-Consumption-Investment Model, 506

6. DIFFERENCE EQUATIONS, 509

- 6.1 Introduction, 509**
- 6.2 Definition and Classification of Difference Equations, 510**
 - Linear Difference Equations, 511
- 6.3 Solutions of Difference Equations, 512**
 - Analogies Between Difference Equations and Differential Equations, 514

- 6.4 Linear First-Order Difference Equations with Constant Coefficients, 515**
 - Behavior of the Solution Sequence, 518
 - Equilibrium and Stability, 522
- 6.5 Applications of Difference Equations in Economic Models, 523**
 - The Harrod Model, 523
 - The General Cobweb Model, 524
 - Consumption Model, 524
 - Income-Consumption-Investment Model, 525
- 6.6 Linear Second-Order Difference Equations with Constant Coefficients, 528**
 - Behavior of the Solution Sequence, 531
 - Nonhomogeneous Second-Order Difference Equations, 533
 - Equilibrium and Stability, 535
- 6.7 Applications of Second-Order Difference Equations in Economic Models, 537**
 - Samuelson's Interaction Model, 537
 - Metzler's Inventory Model, 539

7. MATRIX ALGEBRA, 541

- 7.1 Introduction, 541**
- 7.2 Definition of a Matrix, 541**
 - Definition of a Vector, 542
- 7.3 Matrix Operations, 544**
 - Addition and Subtraction of Matrices, 544
 - Multiplication of a Matrix by a Scalar, 545
 - Multiplication of Matrices, 546
- 7.4 Special Types of Matrices, 551**
 - Diagonal Matrices, 551
 - Identity Matrices, 552
 - Null Matrices, 553
- 7.5 The Transpose of a Matrix, 554**
 - Transpose of a Sum or Difference of Matrices, 556
 - Transpose of a Product of Matrices, 556
- 7.6 Partitioned Matrices, 558**
- 7.7 Determinant of a Matrix, 564**
 - Properties of Determinants, 567
- 7.8 The Inverse of a Matrix, 569**
 - Inversion of 2×2 Matrices, 570
 - Inversion of Larger Matrices, 573
 - Inversion Using Row or Column Operations, 574
 - Steps in Converting a Square Matrix to the Identity Matrix (Gaussian Elimination), 574
 - Inversion Using Adjoints and Determinants, 577
 - Inversion of Partitioned Matrices, 580
 - Properties of Inverses, 582
- 7.9 Simultaneous Linear Equations, 585**
 - Linear Dependence and Rank, 585
 - Solution of Simultaneous Linear Equations, 589
- 7.10 Vector Differentiation, 593**
 - Vector Differentiation of a Linear Function, 593
 - Vector Differentiation of a Vector of Functions, 595
 - Vector Differentiation of a Quadratic Form, 595
 - Vector Differentiation of a Bilinear Form, 597
 - Applications of Vector Differentiation in Maximization and Minimization, 598

8. APPLICATIONS OF MATRIX ALGEBRA, 602

- 8.1 Introduction, 602**
- 8.2 Maxima and Minima of Functions of n Variables, 602**

Unconstrained Maxima and Minima, 602

Lagrange Multipliers, 607

Kuhn-Tucker Conditions, 611

8.3 Input-Output Analysis, 614

8.4 Linear Programming, 619

Geometric Solution, 621

The Simplex Method, 625

The Dual of a Linear Programming Problem, 631

8.5 Game Theory, 639

Number of Persons, 639 The Payoff, 640 Strategies, 640 The Game Matrix, 640 Saddle Points, 641

Solution of Two-Person, Two-Strategy Games, 642

2×2 Games, 643 Solution of 2×2 Games by Matrix Algebra, 646

$2 \times n$ Games and $m \times 2$ Games, 650 Dominance, 651 Graphical Solution, 653

Solution of Larger Games, 656

8.6 First-Order Markov Processes, 661

Definition of a First-Order Markov Process, 661

Steady State or Equilibrium, 662

Algebra Review Problems, 669

Tables, 675

Selected References, 679

Index, 683

Preface to the Third Edition

The general outline of the previous editions is retained in the third edition. Within this general framework, there has been considerable revision. Throughout the book the exposition has, hopefully, been clarified; verbal explanation has been added to many of the examples.

Some new material has been added. Specifically, additional material concerning trigonometric functions is included. Capital formation and continuous compounding are discussed as applications of integration. Several statistical concepts are considered; probability is obtained by integration as the area under a curve, and joint probabilities are obtained by double integration under a surface. The least squares estimates for bivariate linear regression are derived as an application of maximization and minimization using partial differentiation. Vector derivatives of linear functions, quadratic forms, and bilinear forms are considered; the application of vector differentiation in obtaining maxima and minima of functions of several variables is illustrated by the derivation of the least squares estimates for multivariate linear regression. Problems have also been added. A revised solutions manual, including complete solutions to all problems, is available to instructors adopting the text.

Most of the changes, particularly with respect to additional material, were made in response to suggestions by users and reviewers. These suggestions were very helpful and are greatly appreciated; even more of them would have been followed if excessive length were not a consideration.

Every effort has been made, by author and publisher, to avoid errors in the text and solutions manual. The author accepts responsibility for any errors which have gone unnoticed.

J. E. W.

Preface

This text offers the student an understanding of quantitative techniques both in their mathematical contexts and as they are applied to business and economic problems.

The material has been used in various stages of development in graduate and undergraduate quantitative analysis classes at The University of Wisconsin. It assumes reasonable facility with algebra; there is, however, an algebra review section at the end of the book for students whose skills in this area need sharpening. The text is designed for a one-year course, allowing perhaps for some omissions at the option of the instructor.

There are eight chapters: Graphical Representation, Differential Calculus: Functions of One Variable, Differential Calculus: Functions of More Than One Variable, Integral Calculus, Differential Equations, Difference Equations, Vectors and Matrices, and Applications of Matrix Algebra. Each chapter includes mathematical presentation of its subject and applications to business and economics. Each also includes many problems, for we believe that substantial practice in problem solving is essential in the study of mathematics. An extensive solutions manual providing complete calculations for all problems is available on request to instructors who adopt the text.

We wish to express our gratitude to the students who have used this material through its many drafts and whose comments and suggestions have been incorporated in the revisions. We would also like particularly to thank Myrna Edmonds for her patient and painstaking typing of the manuscript. Finally, we accept full responsibility for any errors that may remain.

J. E. D.

J. S. K.

Introduction

From the beginning of recorded history, cultural and scientific advances have depended on the use of symbols. The history of civilization can be viewed as the history of man's increasingly sophisticated use of symbols. As thinking in any area develops, the symbols used become increasingly abstract.

When the concepts to which symbols refer are essentially nonquantitative in nature, the symbols and their relationships can be studied in the general framework of logic and there is no need for mathematics as such. It is when symbols represent essentially quantitative concepts that mathematics is useful and, in fact, essential for analyzing their relationships. Mathematics is a branch of logic, the branch of logic that provides a systematic framework within which quantitative relationships can be studied. In pure mathematics definitions or axioms and assumptions are precisely stated symbolically, and the analysis proceeds by deduction to obtain conclusions. Applied mathematics differs from pure mathematics in a very important respect—in pure mathematics symbols represent abstract concepts whose properties are assigned by definition, while in applied mathematics many symbols represent variables observed in the real world; the properties of these variables must be determined by observation, not by abstract definition, and then stated mathematically. In addition, the empirical accuracy of the deductions of applied mathematics can be determined. Applied mathematical analysis thus is based on empirically determined definitions and assumptions from which empirically verifiable conclusions are obtained by deduction. Pure and applied mathematical analysis differ only with respect to the empirical aspect of the definitions and assumptions and conclusions, not with respect to the methods of deduction.

Because economics is concerned with concepts that are essentially quantitative in nature—for example, price, cost, wage rates, investment, income, and profit—much of economic analysis is inescapably mathematical in nature. Mathematics provides a logical, systematic framework within which quantitative relations can be studied. When economic variables are represented by symbols and their properties stated mathematically, mathematics provides the techniques for analyzing relations among the symbols and thus among the variables they represent. Much of economic analysis is therefore applied mathematical analysis.

In economic analysis, as in applied mathematics in general, deductions obtained by mathematical analysis are interpreted and evaluated empirically. It should be

noted in this regard that if the deductions following from a set of definitions and assumptions are not correct with respect to empirical observation, mathematical analysis (if correctly performed) is not responsible and the difficulty is to be found in the definitions or assumptions. Mathematics enables the economist to be precise in defining relevant variables, to state clearly the assumptions made, to be logical in developing the analysis, and to consider a larger number of variables than might be feasible verbally. However, it does not and cannot prevent the omission or the empirically incorrect definition of relevant variables nor can it prevent empirically inaccurate or incomplete statements of assumptions. Mathematical analysis takes definitions and assumptions as given and obtains the conclusions that follow logically from them. Mathematical analysis is thus by nature logical, not empirical, and can be held responsible for conclusions only with respect to their logical validity given the definitions and assumptions on which they are based and not with respect to their empirical accuracy.

Thus if mathematical analysis is correctly performed, but its conclusions are empirically incorrect, the definitions and assumptions must be examined for accuracy and completeness. By providing a systematic framework for the deduction of empirically verifiable conclusions, mathematical analysis helps the economist determine the accuracy of his definitions and assumptions—if the conclusions are untenable, the definitions and assumptions must be examined and revised.

The purpose of this book is to help the student understand, appreciate, and perform applied mathematical analysis. Mathematical proofs are minimized except when they can be made heuristic. The book is organized so that a type of analysis is discussed first with respect to its mathematical (logical) procedure and then with respect to its applications in economics and business. The assumptions required for each type of analysis are emphasized and its applications are discussed in terms of these assumptions.

The subsequent sections of this chapter review some of the basic mathematical concepts and definitions used in later chapters.

Sets

A *set* is a collection of distinguishable well-defined objects or entities. The objects or entities belonging to a set are said to be *elements* of the set. A set is determined either by a list of its elements or by specifying a rule that determines whether a given object or entity belongs to the set. Such a rule is referred to as a *defining relation*. Braces are used to denote a set; either the elements of the set or its defining relation is written inside the braces.

EXAMPLE

$A = \{a, b, c\}$ means that the set A consists of the elements a , b , and c .

$B = \{x : x \text{ is an odd integer}\}$ means that the set B consists of the odd integers.

$C = \{1, 2, 3, 4, 5, 6\}$ means that the set C consists of the numbers 1, 2, 3, 4, 5, and 6.

$D = \{y : y \text{ is an integer}\}$ means that the set D consists of the integers.

The notation $x \in S$ means that the entity or object x is an element of the set S . The notation $x \notin S$ means that x is not an element of the set S .

A defining relation that is satisfied by no element is said to define the *empty set*, denoted by \emptyset .

EXAMPLE

Referring to the preceding example,

$$a \in A$$

$$b \notin B$$

$$4 \notin B$$

$$4 \in C$$

$$4 \in D$$

$$7 \notin C$$

$$7 \in B$$

$$7 \in D$$

$$d \notin A$$

$$d \notin D$$

EXAMPLE

$$S = \{x : x \text{ is an odd number ending in } 2\} = \emptyset$$

$$P = \{y : y \text{ is an odd number that is the square of an even number}\} = \emptyset.$$

If every element of a set S is also an element of a set T , then S is said to be a *subset* of T . If there is at least one element in T that is not also in S , then S is a *proper subset* of T . The notation $S \subset T$ means that S is a subset of T ; the notation $S \not\subset T$ means that S is not a subset of T . Note that the empty set is a subset of every set. If $S \subset T$ and $T \subset S$, then every element in S is also in T and vice versa and S and T are the same set. This is denoted $S = T$.

EXAMPLE

Referring to the preceding examples,

$$A \not\subset B$$

$$B \subset D$$

$$B \not\subset C$$

$$C \subset D$$

The *union* of two sets S and T is the set consisting of the elements of S and the elements of T . The union of S and T , denoted by $S \cup T$, is described by

$$S \cup T = \{x : x \in S \text{ and/or } x \in T\}$$

Note that an element common to both sets is not counted twice in the union.

The *intersection* of two sets S and T is the set consisting of the elements that are common to S and T . The intersection of S and T is denoted by $S \cap T$ and is described by

$$S \cap T = \{x : x \in S \text{ and } x \in T\}$$

The *set difference* of two sets S and T is the set consisting of the elements of S that are not elements of T . The set difference of S and T is denoted by $S - T$ and is described by

$$S - T = \{x : x \in S \text{ and } x \notin T\}$$

A *universal set* or *universe* is a specified set that contains all the elements of interest for a particular discussion. If two sets have the property that their union is the universal set and their intersection is the null set, one set is said to be the *complement* of the other with respect to the universe. If U denotes the universal set, the complement of a given set S is denoted by S' and is described by

$$S' = U - S$$

A statement of complementation must be qualified unless the universal set is clearly understood. For example, if the complement of S is taken with respect to U and also with respect to V , the notation S'_U and S'_V is used.

EXAMPLE

Referring to the preceding example,

$$B \cup D = D$$

$$A \cup C = \{a, b, c, 1, 2, 3, 4, 5, 6\}$$

$$A \cap D = \emptyset$$

$$B \cap D = B$$

$$B \cap C = \{1, 3, 5\}$$

$$C \cap D = C$$

$$C - D = \emptyset$$

$$C - B = \{2, 4, 6\}$$

$$D - A = D$$

$$D - B = \{x : x \text{ is an even integer}\}$$

If $U = \{x : x \text{ is an integer}\}$, then

$$B'_U = \{x : x \text{ is an even integer}\}$$

$$D'_U = \emptyset$$

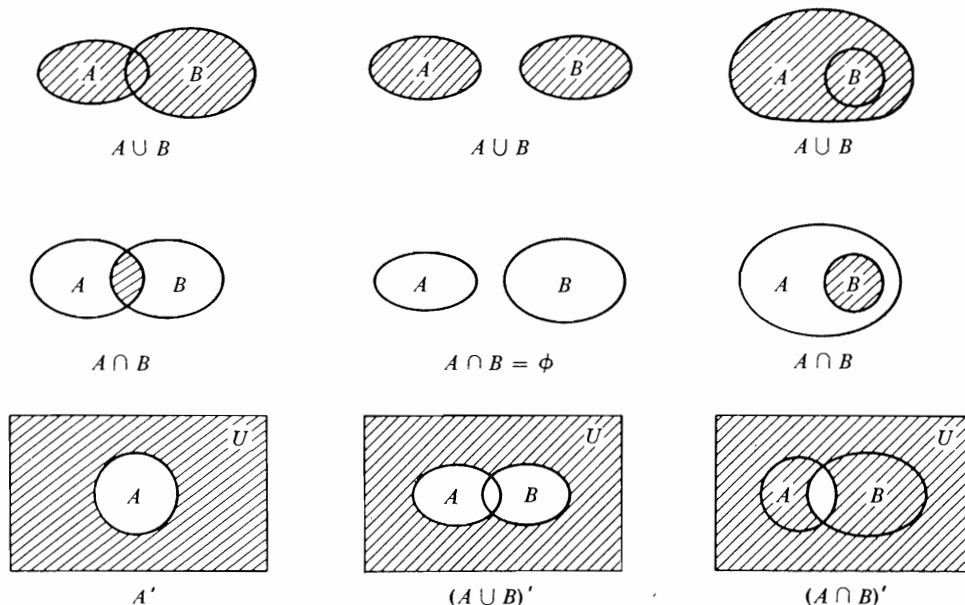
If $V = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$, then

$$C'_V = \{7, 8, 9, 10\}$$

If $W = \{x : x \text{ is a lowercase letter in the English alphabet}\}$, then

$$A'_W = \{d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z\}$$

Sets formed by unions, intersections, and complements are shown here diagrammatically; the shaded area is the set indicated below each diagram. This type of representation of sets is called a *Venn diagram*.



PROBLEMS

- If A , B , and C are sets such that $A \subset B \subset C$, what is the relationship between $C - B$ and $C - A$?
- Show that, in general, $(A \cap B)' = A' \cup B'$.
- Show that, in general, $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.
- Show that $(S \cap T) \cup (S \cap T) = S \cap (S \cup T) = S \cup (S \cap T) = S$.
- Show that $K \cup (L \cap M) = (K \cup L) \cap (K \cup M)$.
- If $A \cap B = \emptyset$ and $A \cap C = \emptyset$, is it necessarily true that $B \cap C = \emptyset$?
- If $A \neq B$ and $B \neq C$, is it necessarily true that $A \neq C$?
- If $A \not\subset B$ and $B \not\subset C$, is it necessarily true that $A \not\subset C$?
- If $A \subset C$ and $B \subset D$, is it necessarily true that $A \cup B \subset C \cup D$?
- If $A \subset C$ and $B \subset D$, is it necessarily true that $A \cap B \subset C \cap D$?
- If $S \cup T = \{1, 2, 3, 4\}$, $S \cap T = \{1, 3\}$, and $S - T = \{2\}$, find S and T .
- If $A \cap B \neq \emptyset$, $A \cap C \neq \emptyset$, and $B \cap C \neq \emptyset$, is it necessarily true that $A \cap B \cap C \neq \emptyset$?
- If U is a universal set, determine which of the following statements are incorrect and correct them by changing the right-hand side of the equation.

a. $B \cup \emptyset = B$	h. $C \cup C = C$
b. $C \cap U = C$	i. $(D')' = U$
c. $A \cup A' = U$	j. $(A - C) \cup C = A - C$
d. $B \cup U = U$	k. $B \cap (B - D) = B \cup D$
e. $D \cap \emptyset = \emptyset$	l. If $A = B'$, then $B = A'$
f. $A \cap A' = A$	m. $(C - D)' = C' - D'$
g. $B \cap B = \emptyset$	n. $(A \cup D) - D = A - D$
- If $A = \{e, f, g\}$ and $B = \{e, h\}$, find

a. $A - B$	c. $A \cap B$
b. $B - A$	d. $A \cup B$
- If $R = \{w, x, y\}$, $S = \{u, v, w\}$, and $T = \{u, v, w, x\}$, and the universal set is $U = \{u, v, w, x, y, z\}$, find

a. $R' \cap T' \cap S'$	e. $(S \cup T) - T'$
b. $(R - S) \cap T$	f. $(R - T) - (S - R)$
c. $(R' - T') \cup S$	g. $(S - R) - [(T - R) \cup (T - S)]$
d. $(R' \cup S)'$	h. $(T - R) \cup S$
- If $A \cap B = \emptyset$ and $A' = C$, is it necessarily true that $B \subset C$?

Answers to Odd-Numbered Problems

1. $C - B \subset C - A$
7. No
9. Yes
11. $S = \{1, 2, 3\}$
 $T = \{1, 3, 4\}$
13. f. $A \cap A' = \emptyset$ j. $(A - C) \cup C = A \cup C$
g. $B \cap B = B$ k. $B \cap (B - D) = B - D$
i. $(D')' = D$ m. $(C - D)' = C' \cup D$
15. a. $\{z\}$ c. S e. T g. \emptyset
b. $\{x\}$ d. $\{w\}$ f. $\{y\}$ h. S

Variables

In mathematics there are two types of quantities, constants and variables. A *constant* is a quantity whose value remains unchanged throughout a particular problem. An *absolute* or *numerical constant* has the same value in all problems; an *arbitrary* or *parametric constant* (or a *parameter*) has the same value throughout any particular problem but may assume different values in different problems. The *absolute* or *numerical value* of a constant a is denoted by $|a|$ and means the magnitude of a regardless of its algebraic sign. Thus $|+a| = |-a| = |a| = a$.

A *variable* is a quantity that assumes various values in a particular problem; the set of values which a variable assumes is its *range*. A variable may be either continuous or discrete. A *continuous variable* is one that may assume any value within a specified interval of real numbers (possibly all real numbers); successive values of a continuous variable thus may differ by infinitesimal amounts. A *discrete variable* is one that may assume only values specified in a countable range. Note that the number of points or values in any specified real interval, say between 5 and 10, is uncountable and no rule can be given for obtaining all possible such values. However, the number of integers or integer values in any specified interval is countable. The number of points on the real line (or on any part of the real line) is said to be a noncountable or nondenumerable infinity. The number of integers on the real line is a countable or denumerable infinity and any set obtained from the integers by a specified rule is also countable or denumerable.

Thus in any problem there may be

1. Absolute constants that always have the same value; such quantities are either numbers or symbols denoting numbers.
2. Parametric constants that have the same value throughout any given problem but may have different values for different problems; such quantities depend on the particular situation represented in the problem.
3. Variables that assume all values meaningful for the problem; such quantities may vary discretely or continuously and may be restricted, for example, to positive values.

EXAMPLE

In the equation of a straight line,

$$\frac{x}{a} + \frac{y}{b} = 1$$

1 is a numerical constant, a and b are parameters, and x and y are variables.

EXAMPLE

In the equation for the area of a circle,

$$A = \pi r^2$$

π is a numerical constant (approximately equal to 3.1416) and A and r are variables.

In pure mathematics, usually the letters at the beginning of the alphabet are used to represent parameters and the letters at the end of the alphabet are used to represent variables. However, in applied mathematics there are many exceptions to this convention, and a variable is frequently represented by the first letter of its name—for example, p for price, q for quantity, c for cost, s for savings, and so forth.

The only values of either constants or variables considered subsequently are *real numbers*. The real number system includes all the familiar types of numbers—rational numbers (positive and negative integers, fractions and decimal fractions) and irrational numbers that are roots of rational numbers and other numbers that cannot be expressed as fractions; it does not include even roots of negative numbers (which are called *imaginary numbers*) or numbers that have both real and imaginary parts (which are called *complex numbers*). The set of all real numbers is denoted by \mathcal{R} .

EXAMPLE

The equation $x^2 + 6 = 0$ has no real solution since $\sqrt{-6}$ is an imaginary number.

EXAMPLE

The equation $x^5 + 32 = 0$ has the real solution $x = -2$ (it also has four complex solutions not of interest for the present discussion).

The quotient a/b of two numbers a and b is a number x such that $a = bx$. Division by zero is not permissible as a result of this definition. Note that if $b = 0$, $a = 0 \cdot x$, which is true only if $a = 0$, in which case x can be any number. Thus the expressions

$$\frac{a}{0} \quad \text{and} \quad \frac{0}{0}$$

are meaningless. To distinguish between $a/0$ and $0/0$, $a/0$ is sometimes said to be “undefined” and $0/0$ is sometimes said to be “indeterminate.” Note that $0/b = 0$ for $b \neq 0$, since $0/b$ is the value of x for which $b \cdot x = 0$ and must be zero for $b \neq 0$.

Relationships involving inequalities are expressed as follows:

The notation $<$ indicates “less than”

The notation $>$ indicates “greater than”

The notation \leq indicates “less than or equal to”

The notation \geq indicates “greater than or equal to”

Relations and Functions

Many problems in mathematics involve sets of ordered pairs of numbers. For example, representation of an object moving in a straight line involves pairs of numbers specifying distance of the object from the origin and the corresponding time in motion. Representation of demand for a given commodity involves pairs of numbers specifying quantity demanded and the corresponding price.

A set of ordered pairs of real numbers is referred to as a *binary relation*. The set of first elements of a binary relation is called the *domain* of the relation; the set of second elements is called the *range* of the relation. For a given set $\{(x, y)\}$, x and y are referred to as *variables*. The set of values variable x takes on is the domain and x is usually called the *independent variable*; the set of values variable y takes on is the range and y is usually called the *dependent variable*. When the number of variables is clear from the context, a binary relation may be referred to simply as a relation.

EXAMPLE

$S_1 = \{(1, 2), (2, 8), (2, 3)\}$ is a binary relation whose domain is $\{1, 2\}$ and whose range is $\{2, 3, 8\}$.

$S_2 = \{(x, y) : x, y \text{ real numbers, } x \leq y\}$ is a binary relation some of whose members are $(2, 2), (3, 4), (5, 5)$, and $(8, 20)$. Note that $(2, 1), (3, 2)$, and $(25, 20)$, for example, are not members of S_2 .

$S_3 = \{(x, y) : y = x^2, x \in \mathcal{R}\}$ is a binary relation. The domain of S_3 is \mathcal{R} and the range is the set of all nonnegative real numbers.

$S_4 = \{(x, y) : y = x^2 \text{ if } 0 \leq x \leq 2, y = 3 - x \text{ if } 2 < x < 3, \text{ and } y = 3 \text{ if } x = 3\}$ is a binary relation whose domain is the set $x : 0 \leq x \leq 3$ and whose range is the set $\{y : 0 \leq y \leq 4\}$.

If a relation is such that to each element of the domain there corresponds one and only one element of the range, the relation is said to be a *function*. Functions constitute a subset of relations—all functions are relations, but not all relations are functions. Note that the relations S_3 and S_4 of the above example are functions, but the relations S_1 and S_2 are not functions. A special notation is used for functions to denote the element of the range corresponding to an element of the domain. If f denotes a function $\{(x, y)\}$, then the number y associated with a given x is denoted by $f(x)$, read “ f of x .” With this notation, the set of pairs defining f may be written as $\{(x, f(x))\}$ where $y = f(x)$.

EXAMPLE

Referring to the preceding example, S_3 and S_4 can be written, respectively, as

$$S_3 = \{(x, f(x)) : f(x) = x^2, x \in \mathcal{R}\}$$

$$S_4 = \{(x, f(x)) : f(x) = x^2 \text{ if } 0 \leq x \leq 2, f(x) = 3 - x \text{ if } 2 < x < 3, \text{ and } f(x) = 3 \text{ if } x = 3\}$$

Other letters, for example, g , F , G , ϕ , and Γ , are also frequently used to indicate a function. An equation such as $g(x) = x^2 + (1/x)$ provides a rule for finding the second member of a pair whose first member is x . Such an equation or formula is said to define the function, although the function is not the formula but is the set of ordered pairs $\{(x, g(x))\}$ or $\{(x, y)\}$. When a value of x is substituted in the formula for a function, the result is said to be the *function value* or the *value of the function* for that value of x .

In the case of a function involving two variables, whenever the value of the independent variable is specified the value of the dependent variable is determined. It is the independent variable whose value is thought of as being assigned arbitrarily (except for nonpermissible values), thus determining also the value of the dependent variable. It is conventional in applied mathematics to represent the independent variable by x and the dependent variable by y .

For most problems in analytic geometry and other branches of pure mathematics, the choice of independent and dependent variables is a matter of convenience and the conventional designation of x and y is concerned only with graphical representation, as discussed below. For example, in considering the equation

$$x - 4y^2 + 2y + 6 = 0$$

it is clearly more convenient to find pairs of points by regarding y as the independent variable and x as the dependent variable as follows:

$$x = 4y^2 - 2y - 6$$

When the variables are purely mathematical and do not represent quantities in a particular context, there is no basis other than convenience for making the choice of independent and dependent variables. However, when the variables do represent quantities in the context of a particular subject matter, the logic of the situation usually determines the choice of independent and dependent variables. For example, quantity produced is thought of as a primary determinant of total cost rather than vice versa. There are exceptions even in these types of problems; for example, price may be thought of as determining quantity demanded, or quantity demanded may be thought of as determining price.

EXAMPLE

If $f(x) = x^2 - x + 2$, then

$$f(z) = z^2 - z + 2$$

$$f(2) = 4 - 2 + 2 = 4$$

$$f(-3) = 9 + 3 + 2 = 14$$

$$f(0) = 0 - 0 + 2 = 2$$

$$f(a) = a^2 - a + 2$$

$$\begin{aligned} f(x+2) &= (x+2)^2 - (x+2) + 2 \\ &= (x^2 + 4x + 4) - (x+2) + 2 \\ &= x^2 + 3x + 4 \end{aligned}$$

$$\begin{aligned} f(x+h) - f(x) &= (x+h)^2 - (x+h) + 2 - (x^2 - x + 2) \\ &= x^2 + 2xh + h^2 - x - h + 2 - x^2 + x - 2 \\ &= 2hx + h^2 - h \end{aligned}$$

EXAMPLE

If $y = f(x) = (x + 1)/x$, then

$$f(1) = \frac{1+1}{1} = 2$$

$$f(-1) = \frac{-1+1}{-1} = 0$$

$$f(0) = \frac{1}{0} \quad (\text{undefined})$$

$$\begin{aligned} f(a+h) - f(a) &= \frac{a+h+1}{a+h} - \frac{a+1}{a} \\ &= \frac{a^2 + ah + a - a^2 - ah - a - h}{a(a+h)} \\ &= -\frac{h}{a(a+h)} \end{aligned}$$

A functional form can be obtained by the substitution of one form into another. If $y = f(x)$ and $u = g(y)$ and if $u = g[f(x)] = h(x)$, then h is called the g composite of f .

EXAMPLE

If $f(x) = x^2 - x - 1$ and $g(x) = x - 1$, then

$$f[g(x)] = (x-1)^2 - (x-1) - 1 = x^2 - 3x + 1$$

and

$$g[f(x)] = (x^2 - x - 1) - 1 = x^2 - x - 2$$

The above example illustrates the fact that, in general, $f[g(x)] \neq g[f(x)]$.

EXAMPLE

If $g(x) = x^2 + 2$, then

$$g[g(x)] = (x^2 + 2)^2 + 2 = x^4 + 4x^2 + 6$$

EXAMPLE

If $f(x) = 10^x$ and $\phi(x) = \log_{10} x$, show that $f[\phi(x)] = \phi[f(x)] = x$.

$$f[\phi(x)] = 10^{\phi(x)} = 10^{\log_{10} x} = x$$

$$\phi[f(x)] = \log_{10} f(x) = \log_{10} 10^x = x$$

PROBLEMS

1. For each of the following relations, state the domain and range and indicate whether the relation is a function.

- $S = \{(1, 3), (2, 3), (2, 4), (3, 2), (4, 1), (5, 5)\}$
- $A = \{(1, 3), (2, 3), (3, 3), (4, 3)\}$
- $T = \{(x, y) : y = 4x + 1 \text{ if } 0 \leq x \leq 2, y = 10 - x^2 \text{ if } 2 < x \leq 3\}$
- $B = \{(x, y) : y^2 = x, y \text{ is an integer and } |y| \leq 8\}$

2. For each of the following, determine if the set $\{(x, y)\}$ of pairs of real numbers formed according to the given rule is a function.

- $y^2 = x$
- $y^3 = x$
- $y^4 = x$
- $x^2 + y = 1$
- $x + y^2 = 1$
- $x^2 + y^2 = 1$
- $y = x^2 + 4$
- $xy = 1$
- $y = \frac{x^2 + 4}{x - 2}$
- $y = \frac{1}{x^2 - 6}$
- $x = \frac{1}{y^2 - y + 2}$
- $\frac{1}{y^2 + 2} = x$

3. If $f(x) = x^3 - x^2 + 6$, find a. $f(0)$; b. $f(-2)$; c. $f(a)$; and d. $f(y^2)$.

4. If $f(x) = \frac{3x^2 - 8}{x - 1}$, find a. $f(3)$; b. $f(-1)$; c. $f(x - 2)$; and d. $f(a - b)$.

5. If $f(y) = \sqrt{\frac{y^2 - 4}{y}}$, find a. $f(-1)$; b. $f(4)$; c. $f(a^2)$; and d. $f(x + 2)$.

6. If $f(y) = 2^y + y$, find a. $f(0)$; b. $f(-1)$; c. $f(5)$; and d. $f(y + b)$.

7. If $f(x) = 3x - x^2$, find a. $f(1)$; b. $f(-2)$; c. $f(a)$; and d. $f(1/h)$.

8. If $g(x) = \frac{x}{x - 3}$, find a. $g(0)$; b. $g(3)$; c. $g(1/x)$; and d. $g(x + b)$.

9. If $h(x) = 4x - x^2$, find a. $h(2) - h(4)$; b. $h(\frac{1}{2}) \cdot h(2)$; c. $h(a + b) - h(c)$;

and d. $\frac{1}{h(a)} + [h(a)]^2$.

10. Give the domain and range of each of the following relations; determine whether each relation is a function and, if not, explain why.

- $y = x^2 + 6$
- $y = 10x - 5$
- $y = \pm \sqrt{4 - 2x^2}$
- $y = -\sqrt{4 - 2x^2}$
- $y = \sqrt{4 - 2x^2}$
- $y = \frac{9}{10x - 5}$
- $y = \frac{25}{x^2}$
- $y = \frac{x^2 + 16}{x + 4}$
- $y^2 = x$
- $2x + y = 3$
- $3x^2 + 2y = 4$
- $4x^2 + 6y^2 = 10$

11. If $f(x) = \frac{x^2}{3} - x$ and $g(t) = \frac{t^2 + 4}{3t}$, find

a. $f(7) - g(3)$ and b. $\frac{f(3)}{g(2) + 1}$.

12. If $q(x) = p(x) + g(x)$ and $p(x) = \frac{x^2 - 7}{3}$, $g(x) = \frac{7}{x^2}$, find $q(2)$.
13. If $h(x) = x^{3/2}$ and $Q(x) = (x^2 + 1)^{-1}$, find $Q[h(x)]$.
14. If $h(y) = e^y$ and $Q(x) = x^2 + 4$, find $Q[h(y)]$.
15. If $h(x) = x^3 + 3x + 6$ and $g(y) = \frac{y}{1+y}$, find $g[h(2)]$.
16. If $f(x) = \frac{8}{x^3}$, $g(x) = x^2$ and $Q(x) = x^3 - 10$, find $Q[f(-2) + g(2)]$.
17. If $g(t) = t^2 + 3$ and $Q(t) = t^{-1}$, find $Q[g(t)]$.
18. If $f(t) = t^3 + a$ and $g(x) = x^{-3}$, find $g[f(t)]$.
19. If $f(t) = e^{t+a}$, $h(t) = e^{b^2t}$ and $g(y) = y^{b^2}$, find $\frac{g[f(t)]}{h(t)}$.
20. If $h(x) = \frac{4}{5} \ln x$ and $g(x) = e^{2x}$, find $h[g(10)]$.

Answers to Odd-Numbered Problems

1. a. domain: $\{1, 2, 3, 4, 5\}$
range: $\{1, 2, 3, 4, 5\}$
not a function
- b. domain: $\{1, 2, 3, 4\}$
range: $\{3\}$
a function
- c. domain: $\{x : 0 \leq x \leq 3\}$
range: $\{y : 1 \leq y \leq 9\}$
a function
- d. domain: $\{0, 1, 4, 9, 16, 25, 36, 49, 64\}$
range: $\{0, \pm 1, \pm 2, \pm 3, \pm 4, \pm 5, \pm 6, \pm 7, \pm 8\}$
not a function
3. a. 6 c. $a^3 - a^2 + 6$
b. -6 d. $y^6 - y^4 + 6$
5. a. $\sqrt{3}$ d. $\sqrt{\frac{x^2 + 4x}{x + 2}}$
b. $\sqrt{3}$
c. $\frac{\sqrt{a^4 - 4}}{a}$
7. a. 2 d. $\frac{3h - 1}{h^2}$
b. -10
c. $3a - a^2$
9. a. 4 d. $\frac{1 + (4a - a^2)^3}{4a - a^2}$
b. 7
c. $4(a + b - c) - a^2 - b^2 + c^2 - 2ab$
11. a. $\frac{71}{9}$, b. 0
13. $(x^3 + 1)^{-1}$
15. $\frac{20}{21}$
17. $(t^2 + 3)^{-1}$
19. e^{ab^2}

INVERSE FUNCTIONS

The domain and range of any relation may be interchanged to form a new relation. Each pair in the new relation is obtained by interchanging the elements of a corresponding pair in the original relation. Two such sets of pairs are said to be *inverse relations*; each relation is said to be the *inverse* of the other. If both relations are functions, they are called *inverse functions*. The inverse relation is a function if and only if the original function is such that to each element of its range there corresponds one and only one element of its domain. The inverse of a function f is denoted by the symbol f^{-1} . In this notation -1 is not an exponent; it means only that f^{-1} is the inverse of f . If f^{-1} is the inverse function of f , then

$$f[f^{-1}(x)] = x \text{ for all } x \text{ in the domain of } f^{-1}$$

and

$$f^{-1}[f(x)] = x \text{ for all } x \text{ in the domain of } f$$

Providing the algebra can be done, the relation $f^{-1}(x)$ can be found by solving $f[f^{-1}(x)] = x$ as though $f^{-1}(x)$ were a variable in the equation. When a domain is not specified, it is assumed to be the set of all real numbers.

EXAMPLE

Let $g = \{(x, y) : y = 2x - 1\}$. Find the inverse relation and determine whether it is a function.

$$y = 2x - 1$$

$$x = \frac{1}{2}(y + 1)$$

which is a function; thus the inverse function is

$$g^{-1} = \{(x, y) : y = \frac{1}{2}(x + 1)\}$$

Since the letters used to indicate the values of the domain and the range are arbitrary, x and y are used in their customary order.

EXAMPLE

Find the inverse of $f = \{(x, y) : y = x^2, x \geq 0\}$ and determine whether it is a function.

$$y = x^2$$

$$x = \sqrt{y}$$

which is a function; thus the inverse function is

$$f^{-1} = \{(x, y) : y = \sqrt{x}\}$$

Note that if f had as its domain the set of all real numbers in the above example, f^{-1} would not be a function since its domain would be the set of all nonnegative real numbers and its range would be the set of all real numbers.

EXAMPLE

If $y = f(x) = (1 - x)/(1 + x)$, find $x = g(y)$. Solving for x ,

$$y = \frac{1 - x}{1 + x}$$

$$y(1 + x) = 1 - x$$

$$y + xy = 1 - x$$

$$y + xy + x - 1 = 0$$

$$x + xy = 1 - y$$

$$x(1 + y) = 1 - y$$

$$x = \frac{1 - y}{1 + y} = g(y)$$

Note that $f(x) = g(x)$ and thus $f(x) = f^{-1}(x)$.

In summary, $y = f(x)$ is a function if y is uniquely determined for every x , that is, if there is one and only one value of y associated with each value of x . If one and only one value of x is associated with each value of y , then there is an inverse function $x = g(y)$; $y = f(x)$ and $x = g(y)$ are inverse functions if and only if $f[g(y)] = y$ and $g[f(x)] = x$. The inverse function of $f(x)$ is frequently denoted $f^{-1}(x)$; then $f[f^{-1}(x)] = f^{-1}[f(x)] = x$. Note that the choice of letter to represent the independent variable in a function is arbitrary; for example, $f(x)$ and $f(y)$ denote the same function. Because it is customary to denote the independent variable by x and the dependent variable by y , if $g(y)$ is the inverse function of $f(x)$, it is frequently denoted by $f^{-1}(x)$.

PROBLEMS

- Determine for each of the following whether the inverse relation is a function; if it is not, alter the range of the given function so that its inverse is a function.
 - $\{(x, y) : y = x^2 + 1\}$
 - $\{(x, y) : y = 4 - x^2\}$
 - $\{(w, z) : z = \sqrt{1 - w^2}\}$
 - $\{(u, v) : v = |u|\}$
- For each of the following functions, find the inverse function $f^{-1}(x)$ and show that $f[f^{-1}(x)] = f^{-1}[f(x)] = x$.
 - $f(x) = 3x + 2$
 - $f(x) = x/(x - 4)$
 - $f(x) = (x - 2)/(x + 2)$
 - $f(x) = (x + 3)/x$
- Find $f[g(x)]$ and $g[f(x)]$ for each of the following.
 - $f(x) = 1/(x - 1)$ and $g(x) = x^2/(x^2 - 1)$
 - $f(x) = x/(4 - x)$ and $g(x) = x/(x - 4)$
 - $f(x) = g(x) = (x + 1)/(x - 1)$
 - $f(x) = \sqrt{x - 1}$ and $g(x) = 1/(x + 1)$
- If $f(x) = (ax + 1)/(bx - 1)$, find values of a and b such that f is its own inverse—that is, $f[f(x)] = x$.
- If $g(h) = he^{1/h}$ and $F\left(\frac{1}{y}\right) = \frac{y^2}{y^2 + 1}$, find
 - $g[F(t)]$
 - $F[g(t)]$
- If $f(x) = x(x + 1)$, show that $f(x + h) - f(x) = h(2x + 1 + h)$.
- If $f(x) = \frac{1}{x}$, show that $f(x + h) - f(x) = \frac{-h}{x^2 + hx}$.

8. If $g(y) = \frac{y}{1-y}$, show that $\frac{1}{2}[g(y) + g(-y)] = g(y^2)$.
9. If $F(z) = \log z$, show that $F(xy) = F(x) + F(y)$.
10. If $\phi(R) = 2^R$, show that $\phi(R+1) = 2\phi(R)$.
11. If $P(x) = \sqrt{x}$, show that $P(x+h) - P(x) = \frac{h}{\sqrt{x+h} + \sqrt{x}}$.
12. If $f(x) = x^2 - 1$ and $g(x) = 2x + 1$, show that $f[g(x)] = 4x(x+1)$.
13. If $f(x) = \frac{x}{1+x}$, show that $f(x) + f(-x) = 2f(-x^2)$.
14. If $g(y) = y^2$ and $h(y) = \frac{y}{1-y}$, show that $h(y^2) = \frac{g(y)}{1-g(y)}$.
15. If $Q(x) = 2 \ln x$ and $f(x) = x^{3/2}$, show that $Q[f(x)] = \frac{3}{2}Q(x)$.
16. If $f(x) = x^2$, show that $f(x-h) - f(x) = f(h) - 2hx$.
17. If $h(x) = x^{1/3}$, $g(x) = (x^9 + x^6)^{1/2}$, and $Q(x) = x(x+1)^{1/2}$, show that $g[h(x)] = Q(x)$.
18. If $f(y) = \frac{y}{1-y}$ and $g(y) = \frac{y}{1+y}$, show that $f(y) - g(y) = 2f(y^2)$.
19. If $f(y) = \frac{1}{1+y^2}$ and $g(y) = \frac{y^2}{1+y^2}$, show that $f(y) + g(y) + \frac{g(y)}{f(y)} = \frac{1}{f(y)}$.
20. If $f(x) = \frac{x+2}{x-2}$, $g(x) = \frac{1+x^2}{x}$, and $h(x) = \frac{1-x}{1+x}$, show that $f[g(x)] = \frac{1}{[h(x)]^2}$.
21. If $f(x) = x-1$ and $g(x) = \frac{1}{x+1}$, show that $f(x^2)g(x) = f(x)$.
22. If $f(y) = \frac{2y}{1+y}$ and $g(y) = \frac{-y}{1-y}$, show that $f(y)g(y) = f(-y^2)$.

Answers to Odd-Numbered Problems

1. a. inverse a function
b. inverse a function if $y \leq 0$ or $y \geq 0$
c. inverse a function
d. inverse a function
3. a. $f[g(x)] = x^2 - 1$; $g[f(x)] = 1/x(2-x)$
b. $f[g(x)] = x/(3x-16)$; $g[f(x)] = x/(5x-16)$
c. $f[g(x)] = g[f(x)] = x$
d. $f[g(x)] = \sqrt{-x/(x+1)}$; $g[f(x)] = (\sqrt{x-1}-1)/(x-2)$
5. a. $\frac{1}{1+t^2} e^{1+t^2}$ b. $\frac{1}{1+t^2} e^{2/t}$

Graphical Representation

1

1.1 Introduction

The French philosopher, Rene Descartes (1596–1650), developed a form of graphic analysis in which algebraic equations are plotted in terms of geometric curves. Thus it is possible to *see* relationships among such variables as price and quantity; direct, indirect, and total cost; savings, investment, and consumption. In addition, much of the theory of calculus can be presented in geometrical terms.

In order to relate algebra and geometry, and thus make possible this dual representation, use is made of a coordinate system which provides a means for locating specific points in a plane or in space. This correspondence may be established in many ways, but the system most commonly used is rectangular coordinates.

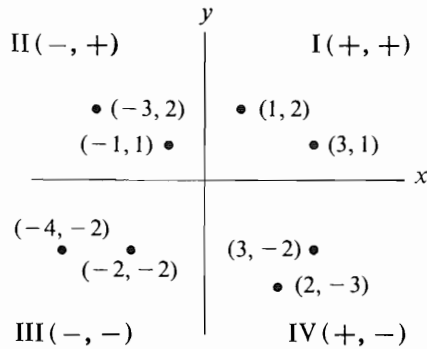
In this chapter, graphical representation using rectangular coordinates is discussed for straight lines and for quadratic, exponential, trigonometric, and logarithmic curves. General methods of graphing and special methods of graphing particular types of curves are considered. Applications of each type of curve in business and economics are also discussed.

1.2 Rectangular Coordinates

For rectangular coordinates, two straight lines intersecting at right angles are used as lines of reference and a point is located in the plane of these lines by giving its perpendicular direction and distance from each of them. These two distances with signs indicating their directions are the *coordinates* of the point. The lines from which the distances are measured are *coordinate axes* or, briefly, *axes*. The point of intersection of the axes is the *origin of coordinates* or, briefly, the *origin*. The coordinate axes divide the plane into four areas or regions called *quadrants*, numbered counterclockwise for reference as shown in Figure 1.1.

Generally the horizontal line is called the *x-axis* and the vertical line is called the *y-axis*. The choice of positive directions is a matter of convenience and may be changed to simplify particular problems. However, it is customary to consider distances measured to the right of the *y-axis* as positive and those measured to the left as

Figure 1.1 The rectangular coordinate system.



negative; similarly, it is customary to consider distances measured upward from the x -axis as positive and those measured downward as negative.

If x and y represent variables having the same physical or geometrical characteristics, or if they are given abstractly as numbers with no physical interpretation, as is frequently the case in analytic geometry, the same unit of measure or scale is used on both axes whenever convenient. If the variables are measured in different units (for example, if x is quantity and y is total cost), the units of measure appropriate for the particular problem are used.

The x -coordinate, or *abscissa*, of a point is that coordinate which indicates the direction and distance of the point to the right or left of the y -axis; the y -coordinate, or *ordinate*, of a point is that coordinate which indicates the direction and distance of the point above or below the x -axis. The position of a point is indicated by writing its coordinates in parentheses in x, y order: (abscissa, ordinate). Locating a point when its coordinates are given is called *plotting* the point.

Fundamental Principle

There is a one-to-one correspondence between number pairs and points in the plane in which they are represented—that is, to each pair of numbers (coordinates) there corresponds one and only one point in the plane and, conversely, to each point in the plane there corresponds one and only one pair of numbers. The terms *point* and *coordinates of a point* are generally used interchangeably.

1.3 Straight Lines

In this section graphical representation of straight lines is considered. The slope of a straight line is defined and four forms of the equation for a straight line are discussed. The relationships among parallel, perpendicular, and intersecting lines are considered and the interpretation of the intersection of two straight lines as the simultaneous solution of the two corresponding equations is discussed. The concept of a family of lines is defined and illustrated.

The rectangular coordinate system can be used to represent any geometric figure

that corresponds to an equation in two variables (for example, an equation in x and y). The geometric figure corresponding to an equation consists of all the points, and only those points, whose coordinates satisfy the equation; in particular, for every straight line in the coordinate plane there is a corresponding linear equation in two variables, and conversely.

An equation of the form

$$Ax + By + C = 0 \quad (1)$$

where A , B , and C are constants and at least one of A and B is nonzero, is said to be *linear in x and y* and such an equation has a straight line as its geometrical representation. Equation (1) is also referred to as the *general equation of the first degree in two variables*.

The *degree of a variable* is the value of the positive integral power to which the variable is raised. The *degree of a term* of an equation in x and y is the sum of its degrees in x and y ; that is, the sum of the powers to which x and y appear in the term. The *degree of an equation* is the degree of its term of greatest degree. For example, $x^3 + y^4 + 6 + y - 12 = 0$ is a fourth-degree equation; $4x^2 + y^4 + 2x^3y^3 - 21 = 0$ is a sixth-degree equation.

There are two aspects of the problem of correspondence between a straight line and a linear equation: (1) given a linear equation, to graph the corresponding line, and (2) given (conditions determining) the line, to find the corresponding linear equation.

The coordinates x and y of every point (x, y) on a given line satisfy the equation corresponding to the line; conversely, the straight line that passes through all the points, and *only* through those points, whose coordinates satisfy the equation is called the *graph* or *locus* of the equation. However, since there are an infinite number of points on any given line or curve, establishing correspondence between algebraic equations and their geometrical representations point by point is clearly not feasible; it is the purpose of analytic geometry to develop methods using a minimum number of points for establishing this correspondence. In general, the more simple the algebraic equation (and thus its corresponding graph), the fewer points necessary for establishing meaningful correspondence.

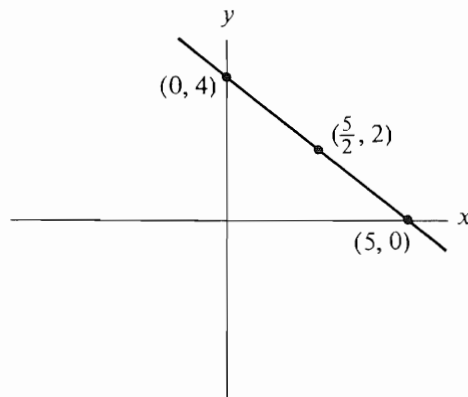
One easy way to graph a straight line is to compute its intercepts. The *intercepts* of a line are the points where the line crosses the axes. Thus the *y-intercept* is the point determined by setting $x = 0$ in the equation of the line; similarly, the *x-intercept* is the point determined by setting $y = 0$ in the equation of the line.

EXAMPLE

Graph the line $4x + 5y - 20 = 0$.

<i>y-intercept</i>	<i>x-intercept</i>
if $x = 0$,	if $y = 0$,
$5y - 20 = 0$	$4x - 20 = 0$
$5y = 20$	$4x = 20$
$y = 4$	$x = 5$
Intercept $(0, 4)$	Intercept $(5, 0)$

Figure 1.2



The intercepts may then be plotted and the line drawn through them, as shown in Figure 1.2. As a check for accuracy, a third point is sometimes plotted as well. For example, let $y = 2$ in the example above. Then

$$4x + 5(2) - 20 = 0$$

$$4x - 10 = 0$$

$$x = \frac{10}{4} = \frac{5}{2}$$

The third point is then $(\frac{5}{2}, 2)$.

SLOPE OF A STRAIGHT LINE

The direction of a straight line is indicated by its slope, which is defined in terms of the angle between the line and the x -axis. When a straight line intersects the x -axis, its angle of inclination is the angle θ , shown in Figure 1.3, which is measured counter-clockwise from the positive direction of the x -axis to the line and thus is always between 0° and 180° . If a line is parallel to the x -axis, its angle of inclination θ is defined to be 0° .

The *slope* of a line is the tangent of its angle of inclination and is usually denoted by m . The definition of the slope of a line is equivalent to the definition of the tangent of its angle of inclination. The tangent is one of the six trigonometric functions of an angle, which are defined in terms of a right triangle formed by the angle and a perpendicular to its adjacent side.

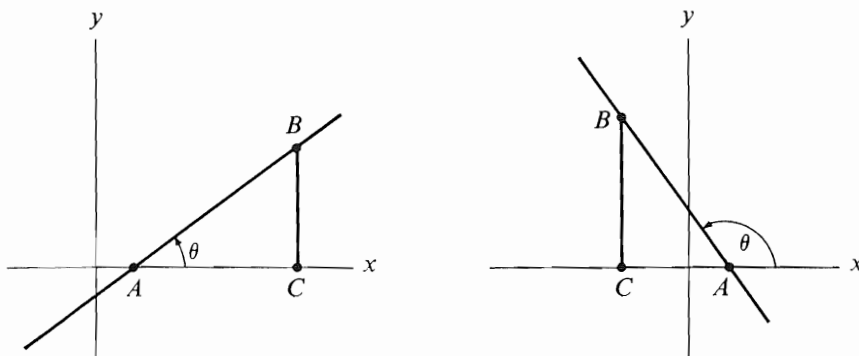
Suppose $0^\circ < \theta < 180^\circ$, and let a line having angle of inclination θ intersect the x -axis at A . From any other point on the line drop a perpendicular cutting the x -axis at C , as in Figure 1.3. Now the tangent of θ , denoted $\tan \theta$, is defined by $\tan \theta = CB/AC$, where CB and AC are directed distances.

Equivalently, if (x_1, y_1) and (x_2, y_2) are any two distinct points on a straight line, then the slope of the line is

$$m = \tan \theta = \frac{y_2 - y_1}{x_2 - x_1} = \frac{\Delta y}{\Delta x}$$

It is customary, especially in calculus, to use the notation Δy , read delta y , to represent the difference $y_2 - y_1$. This symbol Δy denotes the change in y or the difference between two values of y . Do not think of Δy as delta times y . Similarly, Δx denotes $x_2 - x_1$.

Figure 1.3



It can be shown, by an argument based on properties of similar triangles, that the slope of a straight line is the same regardless of the two distinct points chosen for calculating it; correspondingly, the tangent of an angle is the same regardless of the point at which the perpendicular intersects the adjacent side, provided the point is not the intersection of the line and the x -axis.

Since

$$\frac{y_2 - y_1}{x_2 - x_1} = \frac{y_1 - y_2}{x_1 - x_2}$$

the slope of a line may be interpreted as the ratio of the directed change in vertical distance to the corresponding directed change in horizontal distance as a point moves along the line in either direction. In some problems (particularly in physics and engineering), $\Delta y = y_2 - y_1$ is called the *rise* and $\Delta x = x_2 - x_1$ is called the *run*; then the slope is defined as the rise per unit of run, that is, slope = rise/run = $\Delta y/\Delta x$.

If $y_1 = y_2$ and $x_1 \neq x_2$, the line through (x_1, y_1) and (x_2, y_2) is parallel to the x -axis, its angle of inclination θ is zero, and $\tan \theta = \Delta y/\Delta x = 0$.

If $x_1 = x_2$ and $y_1 \neq y_2$, the line through (x_1, y_1) and (x_2, y_2) is parallel to the y -axis, its angle of inclination θ is 90° , and $\tan \theta = \Delta y/\Delta x$ is undefined (see Figure 1.4). (Remember that an expression which involves division by zero is said to be *undefined* or *indeterminate* and no numerical value can meaningfully be assigned to it.)

SUMMARY: SLOPE

The slope of the x -axis is zero.

The slope of the y -axis is undefined.

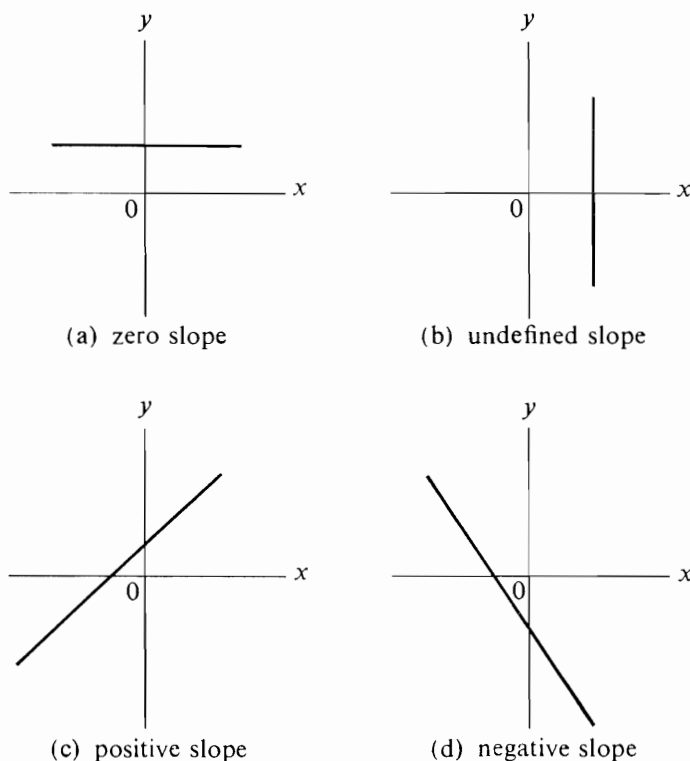
If a line slopes upward to the right, Δx and Δy have the same sign and $\tan \theta = \Delta y/\Delta x > 0$; in this case $0 < \theta < 90^\circ$.

If a line slopes downward to the right, Δx and Δy have opposite signs and $\tan \theta = \Delta y/\Delta x < 0$; in this case $90^\circ < \theta < 180^\circ$.

EQUATIONS FOR STRAIGHT LINES

A straight line, the simplest geometric curve, may be specified uniquely by either of the following: two points which lie on the line or the slope of the line and one point which lies on the line. There are several formulas for obtaining the equation of a straight line; the conditions that are given for obtaining the line determine which formula is most convenient for a particular problem. The two-point form and the intercept form determine a straight line by two points that lie on it; the point-slope form and the slope-intercept form determine a straight line by one point that

Figure 1.4



lies on it and its slope. These formulas are all equivalent to the general equation for a straight line,

$$Ax + By + C = 0$$

and are generally easily transformed into it and into each other.

Two-Point Form

One of the fundamental properties of a straight line is its constant slope; the slope can be determined using any two distinct points on a straight line. These two facts can be used to develop a formula for obtaining the equation of a nonvertical straight line when the coordinates of two points on the line are known.

If (x_1, y_1) and (x_2, y_2) are two distinct points on a nonvertical straight line, then the slope m of the line is given by

$$m = \frac{y_2 - y_1}{x_2 - x_1}$$

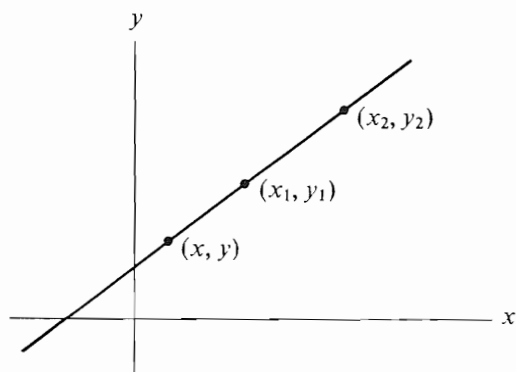
If (x, y) is any other point (that is, a general point) on the straight line, then it and the point (x_1, y_1) can also be used to determine the slope m of the line

$$m = \frac{y - y_1}{x - x_1}$$

and, since the slope is constant,

$$\frac{y - y_1}{x - x_1} = \frac{y_2 - y_1}{x_2 - x_1}$$

Figure 1.5



or

$$y - y_1 = \frac{y_2 - y_1}{x_2 - x_1} (x - x_1) \quad (1a)$$

See Figure 1.5, noting that the labeling of the points is arbitrary and may be chosen for convenience.

Equation (1a), referred to as the *two-point form* for a straight line, is generally most convenient for determining the equation of a straight line when two of its points are given.

EXAMPLE

Find the equation of the line passing through the points (3, 4) and (-5, 2).

$$(x_1, y_1) = (3, 4) \quad (x_2, y_2) = (-5, 2)$$

$$y - y_1 = \frac{y_2 - y_1}{x_2 - x_1} (x - x_1)$$

$$y - 4 = \left(\frac{2 - 4}{-5 - 3} \right) (x - 3)$$

$$y - 4 = \frac{1}{4}(x - 3)$$

$$x - 4y + 13 = 0$$

Check: Insert (3, 4) in $x - 4y + 13 = 0$:

$$3 - 4(4) + 13 = 0$$

$$0 = 0$$

Insert (-5, 2) in $x - 4y + 13 = 0$:

$$-5 - 4(2) + 13 = 0$$

$$0 = 0$$

Intercept Form

For the special case when the point (x_1, y_1) is the y -intercept, denoted by $(0, b)$, where $b \neq 0$, and the point (x_2, y_2) is the x -intercept, denoted by $(a, 0)$, where $a \neq 0$,

then equation (1a) can be written

$$y - b = \frac{-b}{a}(x - 0)$$

$$\frac{y}{b} - 1 = -\frac{x}{a}$$

$$\frac{x}{a} + \frac{y}{b} = 1 \quad (1b)$$

Equation (1b), referred to as the *intercept form* for a straight line, is generally most convenient for determining the equation of a straight line when its two intercepts are given.

EXAMPLE

Find the equation of the line having intercepts $(0, -6)$ and $(4, 0)$.

$$b = -6 \quad a = 4$$

$$\frac{x}{a} + \frac{y}{b} = 1$$

$$\frac{x}{4} + \frac{y}{-6} = 1$$

Multiplying by the lowest common denominator 12,

$$3x - 2y = 12$$

$$3x - 2y - 12 = 0$$

Check: Insert $(0, -6)$ in $3x - 2y - 12 = 0$:

$$3(0) - 2(-6) - 12 = 0$$

$$0 = 0$$

Insert $(4, 0)$ in $3x - 2y - 12 = 0$:

$$3(4) - 2(0) - 12 = 0$$

$$0 = 0$$

Point-Slope Form

Since the slope of a nonvertical line is $m = (y_2 - y_1)/(x_2 - x_1)$, Equation (1a) can be written

$$y - y_1 = m(x - x_1) \quad (1c)$$

Equation (1c), referred to as the *point-slope form* for a straight line, is generally most convenient for determining the equation of a straight line when one of the points on a line and its slope are given.

EXAMPLE

Find the equation of the line passing through the point $(-1, 2)$ and having slope -4 .

$$(x_1, y_1) = (-1, 2) \quad m = -4$$

$$y - y_1 = m(x - x_1)$$

$$y - 2 = -4(x + 1)$$

$$4x + y + 2 = 0$$

Check: $m = -4$ in $4x + y + 2 = 0$. Insert $(-1, 2)$ in $4x + y + 2 = 0$:

$$\begin{aligned} 4(-1) + 2 + 2 &= 0 \\ 0 &= 0 \end{aligned}$$

Slope-Intercept Form

For the special case when the point (x_1, y_1) is the y -intercept, denoted by $(0, b)$, equation (1a) can be written

$$y = mx + b \tag{1d}$$

Equation (1d), referred to as the *slope-intercept form* for a straight line, is generally most convenient for determining the equation of a straight line when its y -intercept and slope are given.

EXAMPLE

Find the equation of the line having y -intercept $(0, 5)$ and slope 3.

$$b = 5 \quad m = 3$$

$$y = mx + b$$

$$= 3x + 5$$

$$3x - y + 5 = 0$$

Check: $m = 3$ in $3x - y + 5 = 0$. Insert $(0, 5)$ in $3x - y + 5 = 0$:

$$\begin{aligned} 3(0) - 5 + 5 &= 0 \\ 0 &= 0 \end{aligned}$$

Vertical Lines

Since the slope of a vertical line is undefined, the above formulas are not appropriate for obtaining the equations of vertical lines. A vertical line passing through the point (x_1, y_1) has the equation $x = x_1$.

EXAMPLE

Find the equation of the vertical line passing through the point $(5, -4)$.

$$x = x_1 = 5$$

$$x = 5$$

Horizontal Lines

Since a horizontal line has zero slope, its equation may be obtained from the two-point, point-slope, or slope-intercept forms. A horizontal line passing through the point (x_1, y_1) has the equation $y = y_1$.

EXAMPLE

Find the equation of the horizontal line passing through the point $(-3, -2)$.

$$y = y_1 = -2$$

$$y = -2$$

SUMMARY: STRAIGHT LINES

General form: $Ax + By + C = 0$

Two-point form: $y - y_1 = \frac{y_2 - y_1}{x_2 - x_1}(x - x_1)$

Intercept form: $\frac{x}{a} + \frac{y}{b} = 1$

Point-slope form: $y - y_1 = m(x - x_1)$

Slope-intercept form: $y = mx + b$

Vertical lines: $x = x_1$

Horizontal lines: $y = y_1$

PROBLEMS

1. a. Which of the following points lie on the line $3x + 4y - 10 = 0$?
 - i. $(1, 2)$
 - ii. $(-2, 4)$
 - iii. $(10, -5)$
 - iv. $(-25, 21)$
 - v. $(0, 0)$
 - vi. $(\frac{22}{3}, \frac{5}{4})$
- b. Graph the line indicating points above which lie on it.
2. For each of the following equations,
 - i. Graph using intercepts.
 - ii. Put into slope-intercept form.
 - iii. Put into intercept form—this form is inappropriate for one of the equations; which one and why?
 - a. $4y - 3x = 12$
 - b. $5x - y = 10$
 - c. $2y + 3x + 2 = 0$
 - d. $x - 3y = 0$
3. a. Which of the following points lie on the line $x - 5y + 4 = 0$?
 - i. $(0, 0)$
 - ii. $(4, 0)$
 - iii. $(1, 1)$
 - iv. $(3, 2)$
 - v. $(0, \frac{4}{5})$
 - vi. $(-1, 5)$
- b. Graph the line, indicating the points above which lie on it.
4. For each of the following pairs of points,
 - i. Find the slope of the line through the two points.
 - ii. Find the equation of the line using the slope.
 - iii. Find the equation of the line without using the slope.
 - iv. Graph the line.
 - a. $(0, 0)$ and $(6, 3)$
 - b. $(\frac{10}{3}, 0)$ and $(0, \frac{5}{3})$
 - c. $(-7, 4)$ and $(8, 4)$
 - d. $(3, -2)$ and $(3, 5)$
 - e. $(-1, -2)$ and $(4, 1)$
 - f. $(-2, -3)$ and $(-5, -6)$

5. Find the equation of the line through the point $(3, -2)$ and perpendicular to the line through the points $(-1, -3)$ and $(3, 7)$.
6. Find the equation of the line through the point $(4, 3)$ and parallel to the line through the points $(0, -3)$ and $(6, 1)$.
7. Find the equation of the line through the point $(5, 15)$ and parallel to the line $y = x + 25$. What relation (parallel, perpendicular, coincident, or intersecting) does this line have to the line through the two points $(6, 0)$ and $(-2, 8)$?
8. Find the equation of the line which has intercept $(0, -3)$ and is perpendicular to the line through the points $(-2, -1)$ and $(2, 5)$.
9. Find the equation of the line which is parallel to the line through the points $(5, 6)$ and $(7, 8)$ and also passes through the intersection of the line having slope -2 through the point $(-4, -6)$ and the line having slope 3 through the point $(2, 2)$.
10. Find the equation of the line passing through the point $(-1, 1)$ and perpendicular to the line having slope $-\frac{1}{4}$ and passing through the point $(5, 2)$.

Answers to Odd-Numbered Problems

1. ii, iii, vi
3. a. $(1, 1)$, $(0, \frac{4}{3})$
5. $2x + 5y + 4 = 0$
7. $x - y + 10 = 0$
perpendicular
9. $x - y - 8 = 0$

PARALLEL, PERPENDICULAR, AND INTERSECTING LINES

Any two lines in a plane are either parallel or intersecting; lines that intersect at right angles are *perpendicular*. Two lines that are *parallel* have equal angles of inclination and therefore have equal slopes, and conversely. Two lines that are perpendicular have slopes which are the negative reciprocals of each other, and conversely.

EXAMPLE

The line $2x + 6y - 4 = 0$ has the indicated relationship to each of the following lines (note that $2x + 6y - 4 = 0$ can be written in the form $y = -\frac{1}{3}x + \frac{2}{3}$):

(a) $4x + 12y - 8 = 0$

$$12y = -4x + 8$$

$$y = -\frac{1}{3}x + \frac{2}{3}$$

coincident with $y = -\frac{1}{3}x + \frac{2}{3}$

(b) $-3x + y - 4 = 0$

$$y = 3x + 4$$

perpendicular to $y = -\frac{1}{3}x + \frac{2}{3}$

(c) $x + 3y - 9 = 0$

$$3y = -x + 9$$

$$y = -\frac{1}{3}x + 3$$

parallel to $y = -\frac{1}{3}x + \frac{2}{3}$

(d) $2x + y - 4 = 0$

$$y = -2x + 4$$

intersecting $y = -\frac{1}{3}x + \frac{2}{3}$

INTERSECTION OF TWO STRAIGHT LINES

The coordinates of the point of intersection of two straight lines must satisfy the equations of both lines. Therefore, the point of intersection of two nonparallel lines can be found by solving their equations simultaneously. The geometric property that two lines intersect corresponds to the algebraic condition that their equations are *independent* and *consistent*, and therefore have a *simultaneous solution*.

Consider two linear equations written in the form

$$A_1x + B_1y + C_1 = 0 \quad A_2x + B_2y + C_2 = 0$$

where A_1 , A_2 , B_1 , B_2 , C_1 , and C_2 are positive or negative constants.

Independent Equations

These equations are *independent* if one cannot be obtained from the other by multiplication by a nonzero constant—that is, they are independent if the equality $A_1/A_2 = B_1/B_2 = C_1/C_2$ does not hold. If the equality does hold, the equations are *dependent*.

NOTE: Two equations that are dependent are equivalent—that is, they represent the same line and have the same graph, since any point that lies on one of the equations lies on the other also.

EXAMPLE

The equations

$$3x + 5y + 10 = 0 \quad 6x + 10y + 20 = 0$$

are dependent or equivalent, since the second equation is obtained by multiplying the first equation by 2; they have the same straight line as their graph.

Consistent Equations

Two equations are said to be *consistent* if they can hold simultaneously—that is, they are consistent if

$$\frac{A_1}{A_2} = \frac{B_1}{B_2} = \frac{C_1}{C_2} \quad \text{or if} \quad \frac{A_1}{A_2} \neq \frac{B_1}{B_2}$$

If $A_1/A_2 = B_1/B_2 = C_1/C_2$, the lines coincide and the equations are both consistent and dependent. If $A_1/A_2 \neq B_1/B_2$, the lines have different slopes and thus intersect at just one point; such lines are consistent and independent. Two linear equations that are *inconsistent*—that is, not consistent—have parallel lines as their graphs.

EXAMPLE

The equations

$$3x + 5y + 10 = 0 \quad 6x + 10y + 15 = 0$$

are inconsistent (since $\frac{3}{6} = \frac{5}{10} \neq \frac{10}{15}$) and cannot both hold simultaneously; they have parallel lines as their graphs.

Note that the equations $A_1x + B_1y + C_1 = 0$ and $A_2x + B_2y + C_2 = 0$ can be written

$$y = \frac{A_1}{B_1}x + \frac{C_1}{B_1}$$

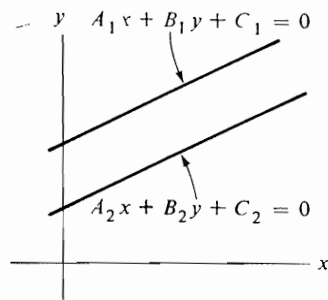
and

$$y = \frac{A_2}{B_2}x + \frac{C_2}{B_2}$$

Thus if $A_1/A_2 = B_1/B_2$, the lines have the same slope. If $B_1/B_2 = C_1/C_2$, the lines have the same intercept. If $A_1/A_2 = B_1/B_2 = C_1/C_2$, the lines have the same slope and the same intercept and are identical.

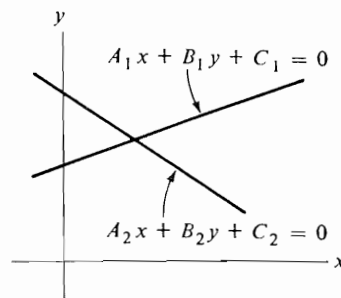
A point is said to be a *simultaneous solution* of two equations if its coordinates satisfy both equations; geometrically, two lines intersect (or coincide) at a point which is their simultaneous solution. Two straight lines have a unique simultaneous solution if they are intersecting (consistent and independent), no simultaneous solution if they are parallel (inconsistent), and infinitely many simultaneous solutions if they are coincident (consistent and dependent). These cases are illustrated in Figure 1.6.

Figure 1.6



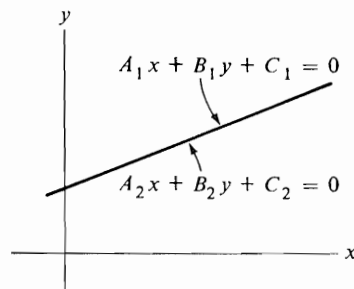
$$(a) \frac{A_1}{A_2} = \frac{B_1}{B_2} \text{ but } \frac{B_1}{B_2} \neq \frac{C_1}{C_2}$$

lines inconsistent and independent



$$(b) \frac{A_1}{A_2} \neq \frac{B_1}{B_2}$$

lines consistent and independent



$$(c) \frac{A_1}{A_2} = \frac{B_1}{B_2} = \frac{C_1}{C_2}$$

lines consistent and dependent

Usually two linear equations can quite easily be solved simultaneously by elimination or substitution; in some cases the method of determinants (to be discussed in detail later) provides the solution more readily. To solve the equations

$$A_1x + B_1y + C_1 = 0 \quad A_2x + B_2y + C_2 = 0$$

by *elimination*, multiply the first equation by A_2 and the second equation by A_1 , subtract to eliminate x , solve for y , and obtain x by substitution in either equation. (Alternatively, multiply the equations by B_2 and B_1 , respectively, and eliminate y .)

To solve by *substitution*, obtain x in terms of y (or, alternatively, y in terms of x) from one equation, substitute in the other equation, and solve.

By either of these methods or, more directly, by the *method of determinants*, the solutions for x and y are

$$x = \frac{B_1C_2 - B_2C_1}{A_1B_2 - A_2B_1} \quad y = \frac{A_2C_1 - A_1C_2}{A_1B_2 - A_2B_1}$$

provided $A_1B_2 - A_2B_1 \neq 0$.

For example, following the procedure for solution by elimination outlined above, multiply the first equation by A_2 and the second equation by A_1 :

$$A_1A_2x + A_2B_1y + A_2C_1 = 0 \quad A_1A_2x + A_1B_2y + A_1C_2 = 0$$

subtract to eliminate x ,

$$(A_2B_1 - A_1B_2)y + A_2C_1 - A_1C_2 = 0$$

solve for y ,

$$y = \frac{A_2C_1 - A_1C_2}{A_1B_2 - A_2B_1}$$

and obtain x by substitution in the first equation

$$\begin{aligned} A_1x + B_1\left(\frac{A_2C_1 - A_1C_2}{A_1B_2 - A_2B_1}\right) + C_1 &= 0 \\ x &= \frac{-B_1}{A_1}\left(\frac{A_2C_1 - A_1C_2}{A_1B_2 - A_2B_1}\right) - \frac{C_1}{A_1} \\ &= \frac{-A_2B_1C_1 + A_1B_1C_2 - A_1B_2C_1 + A_2B_1C_1}{A_1(A_1B_2 - A_2B_1)} \\ &= \frac{B_1C_2 - B_2C_1}{A_1B_2 - A_2B_1} \end{aligned}$$

In many cases it is easier to obtain the simultaneous solution for a particular pair of linear equations directly by elimination or substitution rather than by these formulas.

EXAMPLE

Find the point of intersection of the lines represented by the equations

$$3x - 4y + 6 = 0 \quad x - 2y - 3 = 0$$

(see Figure 1.7)

Eliminating the x -term by multiplying the second equation by -3 and adding it to the first equation,

$$\begin{aligned} 3x - 4y + 6 &= 0 \\ -3x + 6y + 9 &= 0 \\ 2y + 15 &= 0 \\ 2y &= -15 \\ y &= -\frac{15}{2} \end{aligned}$$

Substituting $y = -\frac{15}{2}$ in the first equation,

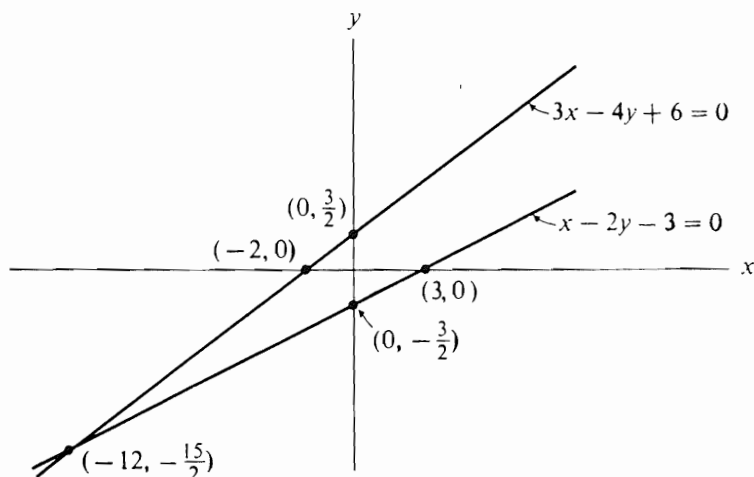
$$\begin{aligned} 3x - 4\left(-\frac{15}{2}\right) + 6 &= 0 \\ 3x + 30 + 6 &= 0 \\ 3x &= -36 \\ x &= -12 \end{aligned}$$

Answer: $(-12, -\frac{15}{2})$.

$$\begin{aligned} \text{Check: Equation 1: } 3(-12) - 4\left(-\frac{15}{2}\right) + 6 &= 0 \\ -36 + 30 + 6 &= 0 \\ 0 &= 0 \end{aligned}$$

$$\begin{aligned} \text{Equation 2: } -12 - 2\left(-\frac{15}{2}\right) - 3 &= 0 \\ -12 + 15 - 3 &= 0 \\ 0 &= 0 \end{aligned}$$

Figure 1.7



Alternatively, substituting for x obtained from the second equation in the first equation:

$$\begin{aligned}x &= 2y + 3 \\3(2y + 3) - 4y + 6 &= 0 \\6y + 9 - 4y + 6 &= 0 \\2y + 15 &= 0 \\2y &= -15 \\y &= -\frac{15}{2}\end{aligned}$$

Substituting for y obtained from the second equation in the first equation:

$$\begin{aligned}y &= \frac{x - 3}{2} \\3x - 4\left(\frac{x - 3}{2}\right) + 6 &= 0 \\3x - 2x + 6 + 6 &= 0 \\x + 12 &= 0 \\x &= -12\end{aligned}$$

Answer: $(-12, -\frac{15}{2})$, as above.

Using the formulas

$$\begin{aligned}x &= \frac{B_1 C_2 - B_2 C_1}{A_1 B_2 - A_2 B_1} & y &= \frac{A_2 C_1 - A_1 C_2}{A_1 B_2 - A_2 B_1} \\A_1 &= 3 & B_1 &= -4 & C_1 &= 6 & A_2 &= 1 & B_2 &= -2 & C_2 &= -3 \\x &= \frac{(-4)(-3) - (-2)(6)}{(3)(-2) - (1)(-4)} \\&= \frac{12 + 12}{-6 + 4} = \frac{24}{-2} = -12 \\y &= \frac{(1)(6) - (3)(-3)}{(3)(-2) - (1)(-4)} \\&= \frac{6 + 9}{-6 + 4} = \frac{15}{-2}\end{aligned}$$

Answer: $(-12, -\frac{15}{2})$, as above.

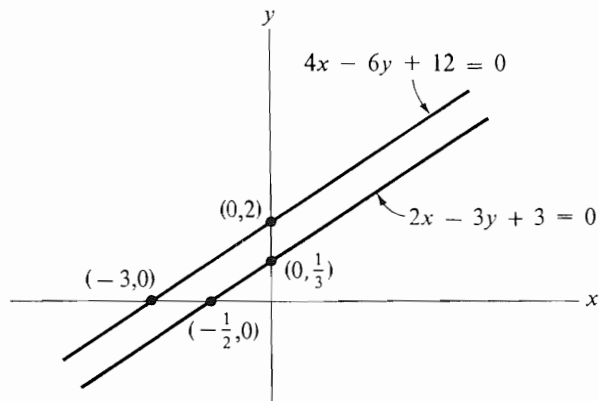
These equations are consistent and independent and have a unique simultaneous solution, the point $(-12, -\frac{15}{2})$.

EXAMPLE

Find the point of intersection of the lines represented by the equations

$$2x - 3y + 3 = 0 \quad 4x - 6y + 12 = 0$$

Figure 1.8



Multiplying the first equation by 2 and rewriting the second equation,

$$4x - 6y + 6 = 0 \quad 4x - 6y + 12 = 0$$

(see Figure 1.8). These equations are inconsistent and do not have a simultaneous solution; they are represented by parallel lines.

EXAMPLE

Find the point of intersection of the lines represented by the equations

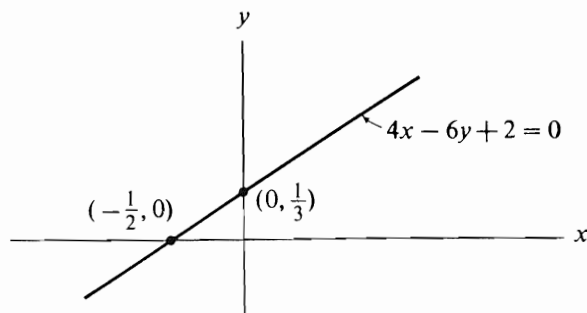
$$2x - 3y + 1 = 0 \quad 4x - 6y + 2 = 0$$

Multiplying the first equation by 2 and rewriting the second equation,

$$4x - 6y + 2 = 0 \quad 4x - 6y + 2 = 0$$

(see Figure 1.9). These equations are dependent and therefore equivalent and have infinitely many simultaneous solutions; they are represented by the same straight line, every point of which is a simultaneous solution of the equations.

Figure 1.9



FAMILIES OF LINES

An important concept in both theoretical and applied mathematics is that of a family of straight lines or curves. The members of a family possess some geometric property (such as passing through a given point) in common, but differ in at least one other geometric property (such as having different slopes). Algebraically, this corresponds to the fact that some of the parameters in a general equation are specified and at least one is not. Remember that an arbitrary (that is, unspecified) constant is called a *parameter*.

The general equation for a straight line

$$Ax + By + C = 0$$

has two parameters, since if any one of the parameters A , B , or C is not zero, it may be used as a divisor to reduce the equation to a form involving only two parameters. The slope-intercept and intercept are two such forms.

The two essential parameters in the equation of a straight line correspond to the two geometric conditions that may be imposed on the line. For each of these geometric conditions there is an algebraic equation showing the corresponding relation between the parameters. Since any given number of variables is determined by that same number of independent linear equations involving the variables, it follows that the number of independent geometric conditions that can be put on a line is the same as the number of essential parameters in the equation of a line, that is, two.

The general equation $Ax + By + C = 0$ with two essential parameters represents all lines in the plane. When one of the two constants is specified and the other is unspecified (that is, remains a parameter) a family of lines—such as all lines of a given slope, or all lines passing through a particular point—is represented by the equation. When both constants are specified, one particular line is represented by the equation.

EXAMPLE

Consider the point-slope form of the equation for a straight line

$$y - y_1 = m(x - x_1)$$

The two essential parameters are expressed in this form of the equation as the slope m and the point (x_1, y_1) . Suppose the condition that the line is to have slope 2 is imposed. This geometric condition corresponds to the algebraic equation

$$y - y_1 = 2(x - x_1)$$

The family of lines represented by this equation consists of an infinite number of lines parallel to each other and all having slope 2. (See Figure 1.10.)

One further condition can be imposed on the line. Geometrically, a point through which the line is to pass can be specified (as a special case, this point could be either the x -intercept or the y -intercept); algebraically, a pair of constant values can be assigned to the coordinates (x_1, y_1) . Suppose it is specified that the line is to pass through the point $(4, 5)$, that is, $(x_1, y_1) = (4, 5)$. (See Figure 1.11.) Now a particular line is determined by the equation

$$y - 5 = 2(x - 4) \quad \text{or} \quad 2x - y - 3 = 0$$

which is obtained from the equations

$$m = 2 \quad \text{and} \quad (x_1, y_1) = (4, 5)$$

corresponding to the imposed geometric conditions; this line has slope 2 and passes through the point $(4, 5)$.

Figure 1.10

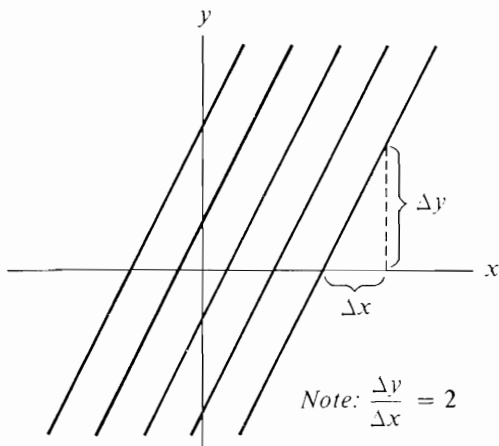
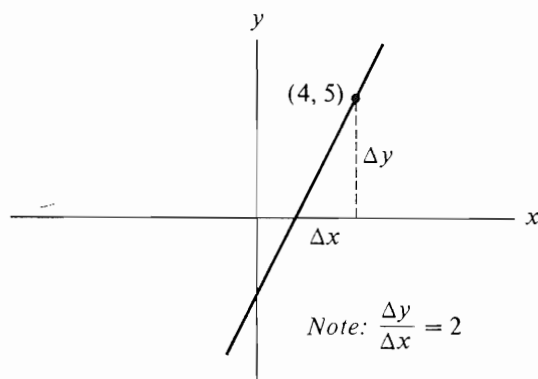


Figure 1.11



EXAMPLE

Consider the intercept form of the equation for a straight line,

$$\frac{x}{a} + \frac{y}{b} = 1$$

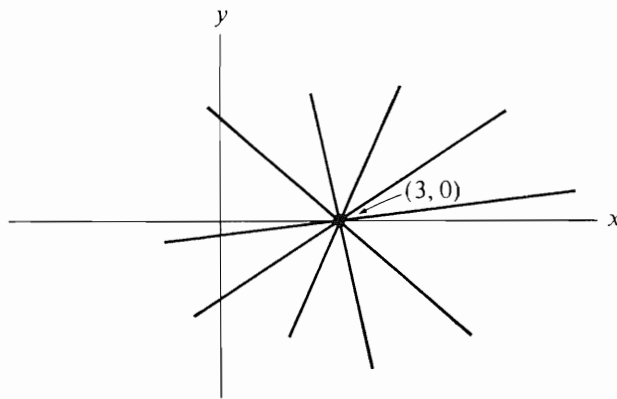
The two essential constants are expressed in this form of the equation as the x -intercept $(a, 0)$ and the y -intercept $(0, b)$. Suppose the condition that the line is to have the x -intercept $(3, 0)$ is imposed. This geometric condition corresponds to the algebraic equation

$$\frac{x}{3} + \frac{y}{b} = 1$$

The family of lines represented by this equation consists of an infinite number of members, each having a different y -intercept (and thus a different slope) but all passing through the point $(3, 0)$. (See Figure 1.12.)

One further condition can be imposed on the line. Geometrically, the y -intercept can be specified; algebraically, a constant value can be assigned to b .

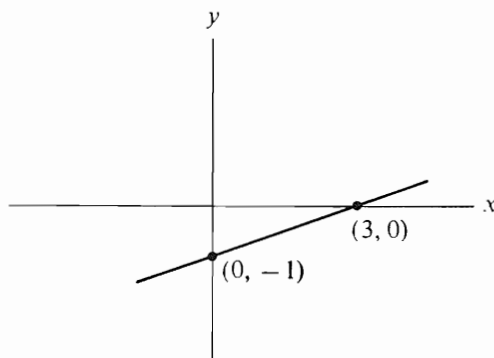
Figure 1.12



Suppose that it is specified that the y -intercept of the line is to be $(0, -1)$, that is, $b = -1$. (See Figure 1.13.) Now a particular line is specified by the equation

$$\frac{x}{3} + \frac{y}{-1} = 1$$

Figure 1.13



or

$$x - 3y = 3$$

$$x - 3y - 3 = 0$$

which is obtained from the two equations

$$a = 3 \quad b = -1$$

corresponding to the imposed geometric conditions; this line has the x -intercept 3 and the y -intercept -1 .

The idea of a general equation representing a large (frequently infinite) class of straight lines or curves and the specification of subclasses of this general class (that is,

families) by imposing certain geometric (and therefore corresponding algebraic) conditions on the lines or curves is very useful in both theoretical and applied mathematics.

PROBLEMS

- What relation (parallel, perpendicular, coincident, or intersecting) does the line $y - 2x - 4 = 0$ have to the following lines?
 - $y - x - 2 = 0$
 - $4y - 8x - 16 = 0$
 - $5y - 10x - 4 = 0$
 - $y - 3x - 4 = 0$
 - $2y + x - 6 = 0$
- What relation (parallel, perpendicular, coincident, or intersecting) does the line $2x - 5y + 6 = 0$ have to each of the following lines?
 - $15x + 6y + 9 = 0$
 - $10x + 4y - 5 = 0$
 - $4x - 10y + 12 = 0$
 - $4x - 8y + 3 = 0$
 - $6x - 15y + 8 = 0$
 - $2x - 5y + 2 = 0$
- What relation (parallel, perpendicular, coincident, or intersecting) does the line $3x + 4y - 2 = 0$ have to each of the following lines?
 - $15x + 20y - 10 = 0$
 - $8x - 6y + 5 = 0$
 - $9x + 12y + 7 = 0$
 - $3x + y - 4 = 0$
 - $12x - 9y + 2 = 0$
 - $2x + y - 6 = 0$
- Determine whether each of the following pairs of equations are a. independent or dependent and b. consistent or inconsistent.
 - $2x - 6y + 5 = 0$ and $3x - 8y + 3 = 0$
 - $x + 5y - 2 = 0$ and $x + 5y - 5 = 0$
 - $3x - 9y + 12 = 0$ and $x - 3y + 4 = 0$
 - $5x - 4y - 6 = 0$ and $4x - 5y + 6 = 0$
- Referring to the pairs of equations in Problem 4, a. determine which pairs of equations have simultaneous solutions and obtain those solutions and b. graph the pairs of equations.
- If a pair of equations are consistent and independent, do they necessarily have a unique simultaneous solution?
- If a pair of equations are consistent and dependent, do they necessarily have a simultaneous solution? If so, is it unique?
- Can a pair of equations be inconsistent and dependent?
- If a pair of equations represent perpendicular lines, are the equations consistent? Are they independent?
- Graph the family of straight lines parallel to the x -axis and give the equation corresponding to this specification. Graph the member of this family that passes through the point $(10, -6)$.
- Graph the family of straight lines passing through the point $(-1, 6)$ and give the equation corresponding to this specification. Graph the member of this family which is parallel to the line $y + 6x - 5 = 0$ and write the equation for this line.
- Graph the family of straight lines perpendicular to the line $2x - 5y - 10 = 0$ and give the equation corresponding to this specification. Graph the member of this family which passes through the point $(4, -1)$ and write the equation for this line.

Answers to Odd-Numbered Problems

- intersecting
 - coincident
 - parallel
 - intersecting
 - perpendicular
- coincident
 - perpendicular
 - parallel
 - intersecting
 - perpendicular
 - intersecting
- i. $(11, \frac{9}{2})$
 - iii. all points on the line $x - 3y + 4 = 0$
 - iv. $(6, 6)$

7. yes; no
 9. yes; yes
 11. $y = mx + 6 + m$; $y = -6x$

1.4 Applications of Straight Lines in Business and Economics

In this section several applications of straight lines in business and economics are discussed. These applications include linear supply and demand curves and the corresponding market equilibria, break-even analysis, and the consumption function.

LINEAR DEMAND AND SUPPLY CURVES

In practice, some demand and supply equations are approximately linear for the relevant range; others are not linear. Even in the latter cases, linear equations may provide reasonably accurate representations of supply and demand for a limited range. In this section linear demand and supply equations are used for simplicity and clarity in illustrating certain types of analyses. Nonlinear equations are discussed in subsequent sections.

Figure 1.14(a) shows a more general representation of supply and demand curves. Figure 1.14(b) represents supply and demand as linear functions. It should be noted, as indicated in Figure 1.14, that only the segments of the equations which fall in the first quadrant are pertinent to economic analysis. This is because supply, price, and quantity are, in general, either zero or positive. For example, in simpler forms of economic analysis:

Negative supply implies that goods are not available in the market either because they are not produced or because they are withheld until a satisfactory price is offered.

Negative price implies that prices are paid to buyers for the removal of goods from the market.

Negative demand implies that price is so high as to preclude market activity until quantities are offered at a satisfactory price.

These cases can occur, but their incidence is infrequent and is considered only in more advanced economic analysis.

It is very important to realize that the straight line, mathematically, is perfectly

Figure 1.14

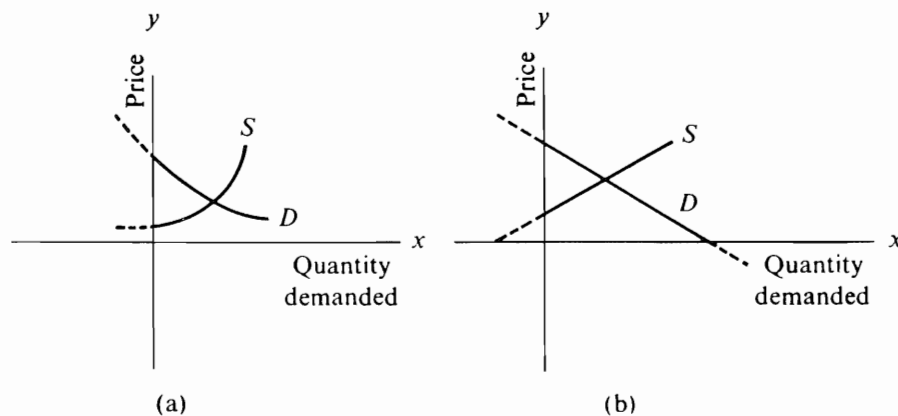
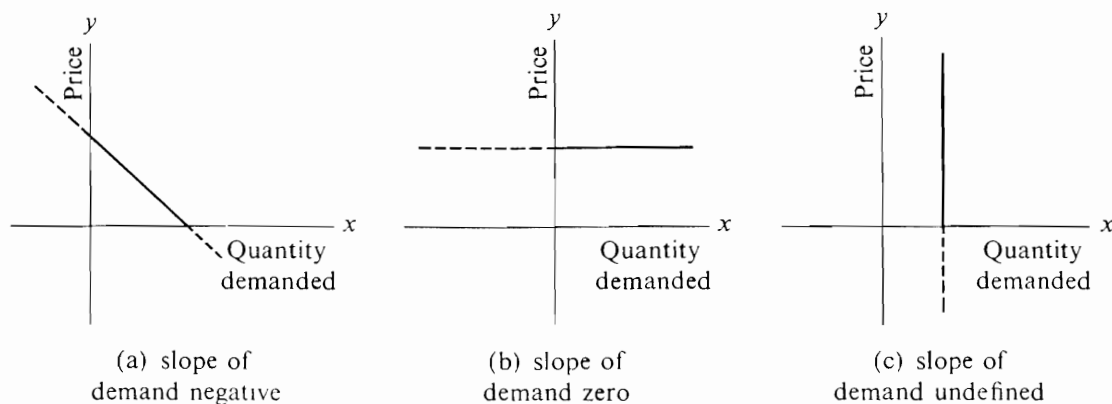


Figure 1.15



general; the formulas for a straight line do not indicate the range of values of x and y which are to be considered. When it is specified, as it is for the present purpose, that only zero or positive values of x and y are of interest—and conversely, that negative prices or quantities are not meaningful—the range of values for x and y is restricted. These restrictions are based on the interpretation and meaning of the equation for a *particular application*; they are *not* based on its inherent mathematical properties. This must be kept in mind to avoid misinterpretations, particularly when more complicated equations are considered.

Linear Demand Curves

In the usual case, the slope of a demand curve is negative—that is, as price increases, quantity demanded decreases, and as price decreases, quantity demanded increases. In certain cases the slope of a demand curve may be zero—constant price regardless of demand. In other cases the slope of a demand curve may be undefined—constant demand regardless of price. Figure 1.15 illustrates these three cases.

Depending on the information available, various formulas for straight lines may be most convenient for obtaining the demand function.

EXAMPLE

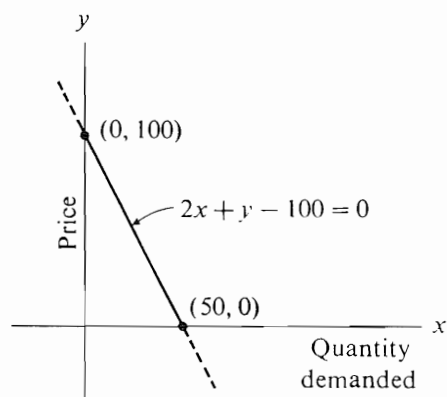
Ten watches are sold when the price is \$80; 20 watches are sold when the price is \$60. What is the demand equation?

$$\begin{aligned}
 y - y_1 &= \frac{y_2 - y_1}{x_2 - x_1} (x - x_1) \\
 x_1 &= 10 & y_1 &= 80 \\
 x_2 &= 20 & y_2 &= 60 \\
 y - 80 &= \left(\frac{60 - 80}{20 - 10} \right) (x - 10) \\
 &= -2(x - 10)
 \end{aligned}$$

$$2x + y - 100 = 0$$

(See Figure 1.16.)

Figure 1.16



EXAMPLE

When the price is \$100, no watches are sold; when watches are free, 50 are demanded. What is the demand equation?

$$\frac{x}{a} + \frac{y}{b} = 1$$

$$a = 50 \quad b = 100$$

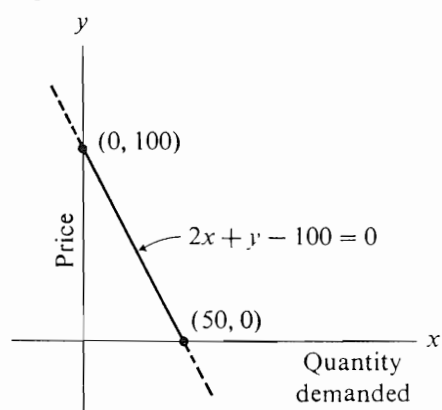
$$\frac{x}{50} + \frac{y}{100} = 1$$

$$2x + y = 100$$

$$2x + y - 100 = 0$$

(See Figure 1.17.)

Figure 1.17



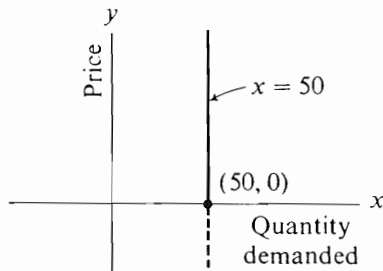
EXAMPLE

Because they are considered necessary for national security, 50 heavy-duty generators are bought every year, regardless of price. What is the demand equation?

$$x = x_1 = 50$$

(See Figure 1.18.)

Figure 1.18



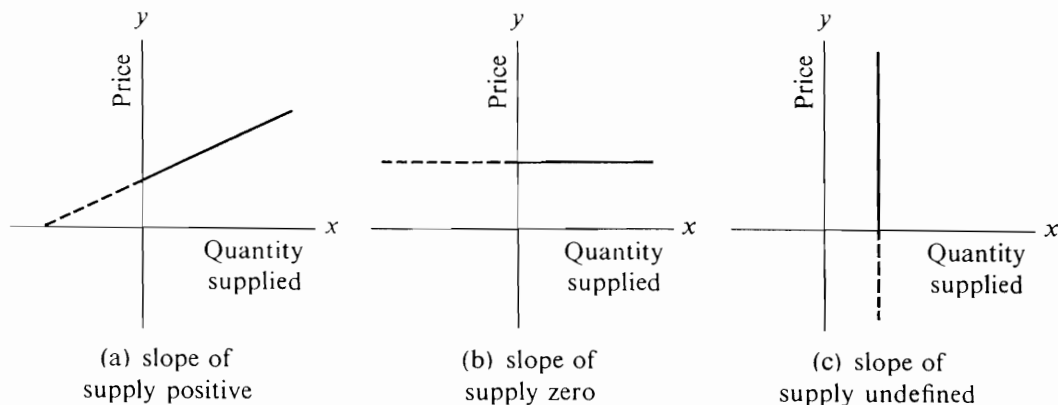
Linear Supply Curves

In the usual case, the slope of the supply curve is positive—that is, as the price increases, quantity supplied increases and as price decreases, quantity supplied decreases. In certain cases, the slope of a supply curve may be zero—constant price regardless of supply. In other cases, the slope of a supply curve may be undefined—constant supply regardless of price (see Figure 1.19).

As in the discussion of demand curves, y represents the price, in appropriate units, and x represents the quantity supplied, in appropriate units. Only positive values of x and y are of interest, as previously stated. Note that for the supply curve in case (a) the y -coordinate of the y -intercept may be positive, negative, or zero.

The x -coordinate of the x -intercept may be negative and, therefore, outside the range of values of interest. This is reasonable since producers usually cease to supply a commodity before the price reaches zero.

Figure 1.19



EXAMPLE

When the price is \$50, 50 cameras of a fixed type are available for market; when the price is \$75, 100 of the cameras are available for market. What is the supply equation?

$$y - y_1 = \frac{y_2 - y_1}{x_2 - x_1} (x - x_1)$$

$$x_1 = 50 \quad y_1 = 50$$

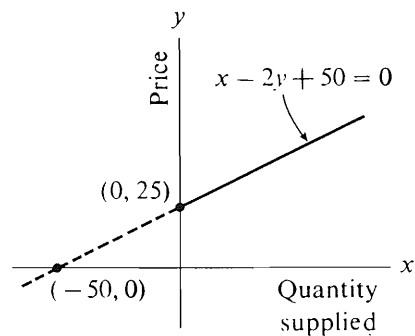
$$x_2 = 100 \quad y_2 = 75$$

$$\begin{aligned} y - 50 &= \left(\frac{75 - 50}{100 - 50} \right) (x - 50) \\ &= \frac{1}{2} (x - 50) \end{aligned}$$

$$x - 2y + 50 = 0$$

(See Figure 1.20.)

Figure 1.20



EXAMPLE

When the price is \$25, no cameras of a fixed type are available for market; for every \$10 increase in price, 20 more cameras are available for market. What is the supply equation?

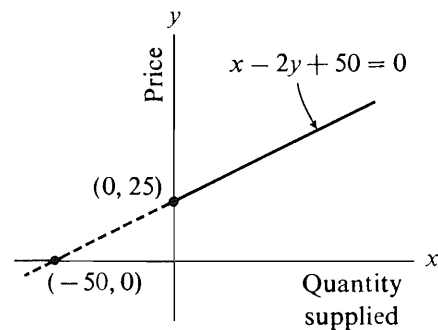
$$y = mx + b \quad m = \frac{1}{2} \quad b = 25$$

$$y = \frac{1}{2}x + 25$$

$$x - 2y + 50 = 0$$

(See Figure 1.21.)

Figure 1.21



EXAMPLE

According to the terms of the contract between Company A and the telephone company, Company A pays the telephone company \$500 per month for long-distance calls of unlimited length of time. What is the supply equation?

$$y = y_1 = 500$$

(See Figure 1.22.)

Figure 1.22

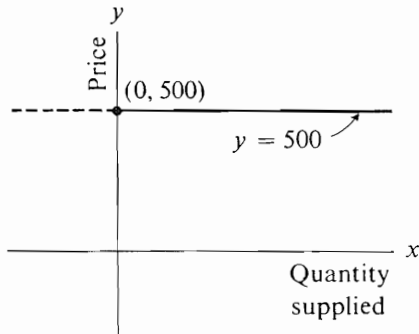
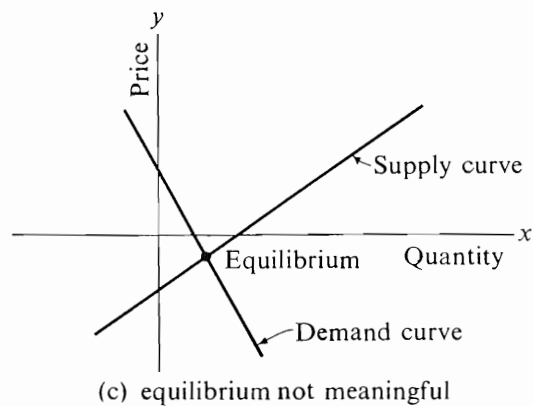
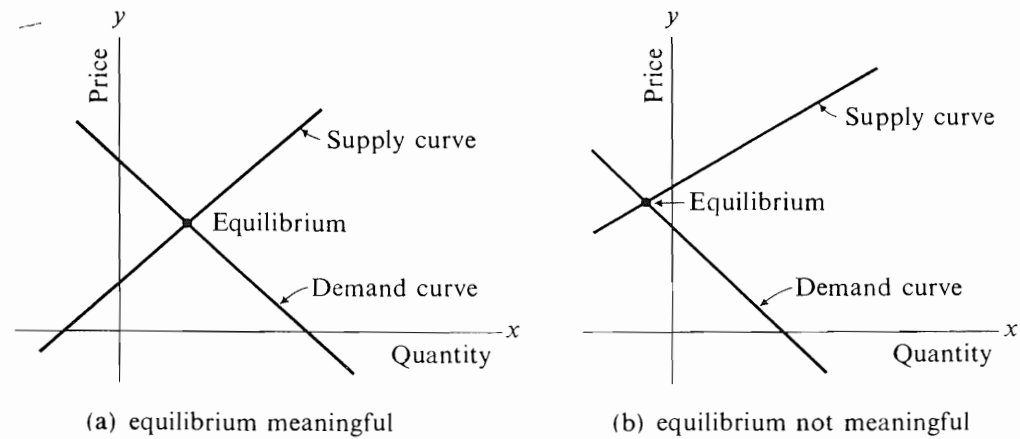


Figure 1.23



Market Equilibrium

Market equilibrium is said to occur at that point (price) at which the quantity of a commodity demanded is equal to the quantity supplied, as shown in Figure 1.23. Thus, provided the same units for x and y are used in both equations, the equilibrium amount and the equilibrium price correspond to the coordinates of the point of intersection of the demand and supply curves. Algebraically, the equilibrium quantity and price are found by solving the supply and demand equations simultaneously (again, provided the same units for x and y are used in both equations).

In general, for an equilibrium to be meaningful, the values of both x and y must be positive or zero—that is, the demand and supply curves must intersect in the first quadrant. This occurs if and only if the y -coordinate of the y -intercept of the demand curve is greater than or equal to the y -coordinate of the y -intercept of the supply curve and the x -coordinate of the x -intercept of the demand curve is greater than or equal to the x -coordinate of the x -intercept of the supply curve. This can be seen geometrically in Figure 1.23; the proof is given in Technical Note I.

EXAMPLE

Find the point of equilibrium for the following demand and supply equations.

$$y = 10 - 2x$$

$$y = \frac{3}{2}x + 1$$

Solving the equations simultaneously by substitution,

$$10 - 2x = \frac{3}{2}x + 1$$

$$\frac{7}{2}x = 9$$

$$x = \frac{18}{7}$$

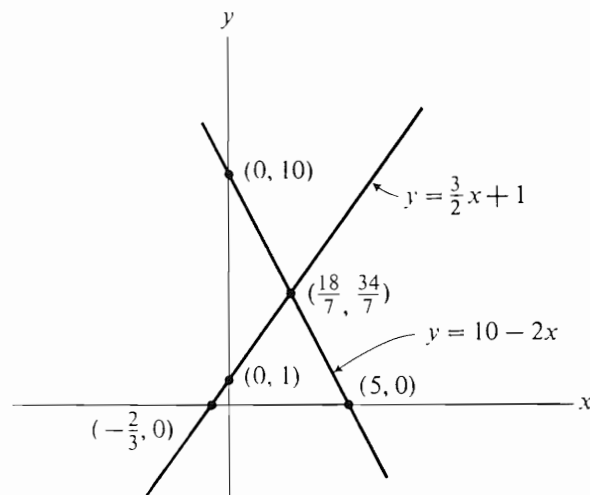
$$y = 10 - 2\left(\frac{18}{7}\right)$$

$$= \frac{34}{7}$$

$$\text{Answer: } \left(\frac{18}{7}, \frac{34}{7}\right).$$

(See Figure 1.24.)

Figure 1.24



EXAMPLE

Find the point of equilibrium for the following demand and supply equations.

$$y = 5 - 3x$$

$$y = 4x + 12$$

Solving the equations simultaneously by substitution,

$$5 - 3x = 4x + 12$$

$$7x = -7$$

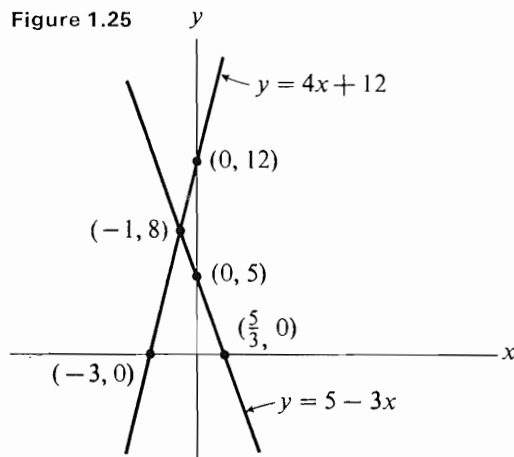
$$x = -1$$

$$y = 5 - 3(-1)$$

$$= 8$$

Answer: $(-1, 8)$ is not a meaningful equilibrium—note that $b_D = 5 < b_S = 12$.
(See Figure 1.25.)

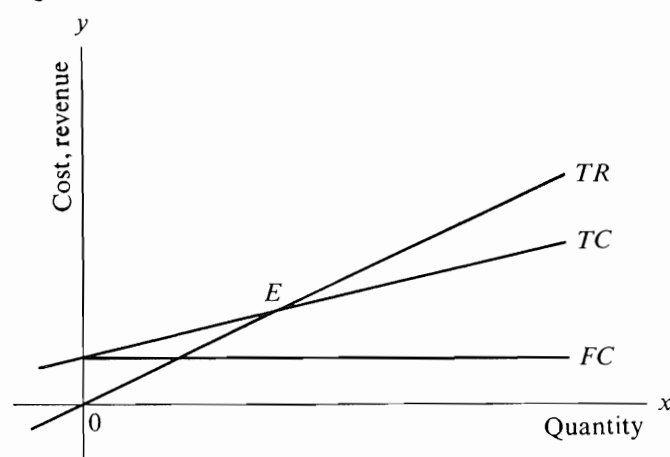
Figure 1.25



BREAK-EVEN ANALYSIS

Break-even graphs are frequently used in business and economics to analyze the implications of various pricing and production decisions. A break-even graph, in simplified form, is shown in Figure 1.26.

Figure 1.26



In this particular case, costs have been divided into two general categories: fixed and variable. Fixed costs remain constant at all levels of output and commonly include such items as rent, depreciation, interest, plant, and equipment; variable costs are those which vary with output and include such items as labor, materials, and promotional expense. Total cost at any level of output is the sum of fixed cost and the variable cost at that level of output.

In Figure 1.26 the straight line FC represents fixed cost. It is the straight line with the y -intercept equal to the constant fixed cost (that is, cost when output is zero) and slope zero.

The line TC represents total cost; its y -intercept is equal to fixed cost and its slope is equal to the increase in variable cost per unit increase in output. Note the assumption that variable cost is proportional to output over the relevant range.

The line TR represents total revenue to the firm for different quantities sold; its intercept is at the origin and its slope is equal to the price per unit, assuming that this price is constant for all quantities sold.

The break-even point E is that point at which the lines TR and TC intersect. It represents the quantity at which the producer just breaks even—that is, the quantity for which there is just sufficient revenue to cover costs. Break-even analysis is used more frequently in practice to demonstrate probable effects of changes than to determine what those changes should be.

EXAMPLE

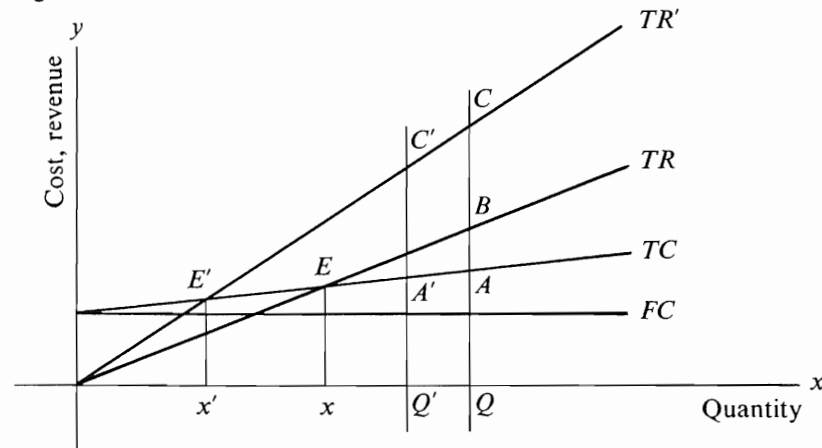
In Figure 1.27, suppose the firm decides that it could sell the same quantity of product, denoted by point Q , if the price per unit were increased so that total revenue would be raised from TR to TR' . TC remains the same, and the break-even point changes from E to E' .

At the original price, profit to the firm is represented by the area of triangle EAB ; at the higher price, it is represented by the area of triangle $E'AC$.

If, on the other hand, the firm assumes that the increase in price resulting in TR' would reduce the quantity sold to that denoted by point Q' , the following would be indicated:

- Total cost reduced from that denoted by A to A'
- Total revenue reduced from that of point C to C'
- Total profit changed from area EAB to area $E'A'C'$

Figure 1.27



Since the profit area $E'A'C'$ is greater than the profit area EAB resulting from the larger quantity sold at the original price, the company might decide to increase the price even though sales would be expected to drop.

Break-even analysis may be used to study possible results of many combinations of factors involved in a particular problem.

EXAMPLE

Suppose the fixed cost of production for a commodity is \$5000; the variable cost is \$7.50 per unit and the commodity sells for \$10 per unit. What is the break-even quantity?

$$5000 = (10 - 7.50)x$$

$$5000 = 2.5x$$

$$x = 2000$$

The break-even quantity is 2000 units.

THE CONSUMPTION FUNCTION

National income analysis provides an interesting example of the use of linear functions, since the consumption function is frequently assumed to be linear over relatively short ranges or "in the short run." The analysis, in one of its simplest forms, is based on the following assumptions:

1. There is some absolute amount of consumption necessary to maintain life, even though there may be no money income.
2. Consumption is related to disposable income, that is, $c = f(y_d)$.
3. When disposable income increases, consumption will also increase, but by a smaller amount. Stated mathematically, if

Δy_d represents an increase in disposable income

and

Δc represents the resulting change in consumption

then

$\frac{\Delta c}{\Delta y_d}$ will be positive but less than one, that is, $0 < \frac{\Delta c}{\Delta y_d} < 1$

4. The proportion of an increment in disposable income which will be consumed is constant. This proportion is referred to as the "marginal propensity to consume."

These assumptions can be translated into the point-slope form of the equation for a straight line

$$c = a + by_d$$

where c represents consumption; a is fixed, basic consumption regardless of income; b is the marginal propensity to consume; and y_d is disposable income.

EXAMPLE

When national disposable income is 0, national consumption is 5 (in billions of dollars). For the economy as a whole consumption is linearly related to national disposable income as follows: At each level of disposable income, consumption equals 5 (in billions of dollars) plus 80 percent of disposable income.

What is the equation that expresses this relationship?

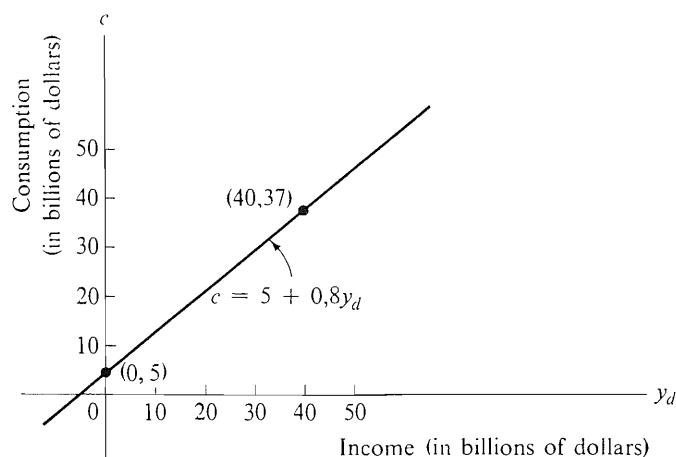
$$c = 5 + 0.8y_d$$

What is aggregate consumption when disposable income is 40 (in billions of dollars)?

$$c = 5 + 0.8(40) = 37$$

Graph the line representing aggregate consumption as a function of national disposable income. (See Figure 1.28.)

Figure 1.28



EXAMPLE

Suppose that the consumption function is given by

$$c = 10 + 0.55y_d$$

where c is aggregate consumption and y_d is disposable income (in billions of dollars). What is aggregate consumption when disposable income is 10 billion dollars? When disposable income is 5 billion dollars, what proportion of it is consumed? When disposable income is 5 billion dollars, what proportion of aggregate consumption is consumption of disposable income?

If $y_d = 10$,

$$\begin{aligned} c &= 10 + (0.55)(10) \\ &= 15.5 \end{aligned}$$

Thus aggregate consumption is 15.5 billion dollars when disposable income is 10 billion dollars.

The proportion of disposable income that is consumed is 0.55 for any level of disposable income.

If $y_d = 5$,

$$\begin{aligned} c &= 10 + (0.55)(5) \\ &= 12.75 \end{aligned}$$

Thus, if disposable income is 10 billion dollars, consumption of disposable income is 2.75 billion dollars, which is $\frac{1}{4}$ or approximately 0.22 of aggregate consumption.

PROBLEMS

- Which of the following equations represent demand curves, which represent supply curves, and which represent neither demand nor supply curves? (Assume y represents price and x represents quantity.)

a. $x - 2y = 0$	f. $2x + 5y + 4 = 0$
b. $3x + 4y - 10 = 0$	g. $3x + 4y - 12 = 0$
c. $y - 4 = 0$	h. $5x - y - 10 = 0$
d. $x - 3 = 0$	i. $2y + 3x + 2 = 0$
e. $2x - 3y + 1 = 0$	j. $x - 3y = 0$
- The demand curve for a commodity is $x = 10 - y/4$. (Assume y represents price and x represents quantity demanded.)
 - Find the quantity demanded if the price is i. 4, ii. 16, iii. 25.
 - Find the price if the quantity demanded is i. 9, ii. 7, iii. 2.
 - What is the highest price that would be paid for this commodity?
 - What quantity would be demanded if the commodity were free?
 - Graph the curve.
- The supply curve of a commodity is $x = 1.1y - 0.1$. (Assume y represents price and x represents quantity supplied.)
 - Find the price if the quantity supplied is i. 1, ii. 0.8, iii. 0.5.
 - Find the quantity supplied if the price is i. 8, ii. 6, iii. 4.1.
 - What is the lowest price at which this commodity would be supplied?
 - Graph the curve.
- The demand equation of a commodity is $x = A - By$, where A and B are positive constants, y represents price, and x represents quantity demanded.
 - Find the price if the quantity demanded is $A/3$.
 - Find the quantity demanded if the price is $A/2B$.
 - Find the quantity demanded if the commodity is free.
 - What is the highest price that would be paid for this commodity?
- The supply equation of a commodity is $x = ay - b$, where a and b are positive constants, y represents price, and x represents quantity supplied.
 - Find the price if the quantity supplied is i. $5a - b$, ii. $a + 2b$.
 - Find the quantity supplied if the price is i. $3b/a$, ii. $5b/a$.
 - What is the lowest price at which this commodity would be supplied?
- For each of the following pairs of straight lines,
 - Determine which is the demand and which is the supply curve.
 - Graph the curves and estimate the market equilibrium price and quantity.
 - Solve the equations algebraically and verify the estimates for market equilibrium price and quantity.

- a. $y = 10 - 2x$ and $y = \frac{3}{2}x + 1$ c. $x = 15 - 3y$ and $x = 2y - 3$
 b. $y = 6$ and $x = 3y - 3$ d. $2y + 3x = 10$ and $x = 4y - 6$
7. A manufacturer sells his product at \$5 per unit.
 - a. What is the total revenue for sales of 5000 units of product? What is the equation for this revenue function? Graph the function.
 - b. Fixed costs are constant at \$3000 regardless of the number of units of product involved. Superimpose the graph of this function on graph a. above.
 - c. Total cost is equal to the sum of fixed costs and variable costs. In this company, variable costs are estimated at 40 percent of total revenue. What is the total cost when 5000 units of product are sold? Graph the function superimposed on graph a.
 - d. What is the break-even point? Indicate this point on the graph, and solve for the corresponding amount sold. Indicate on the graph the quantity at which the manufacturer will cover his fixed costs.
 8. At a price of \$5 per unit, a firm will supply 5000 plastic battery lanterns every month; at \$3.50 per unit it will supply 2000 units. Determine the equation of the supply function for this product. Graph the equation.
 9. For the economy as a whole, consumption is linearly related to national disposable income as follows: at each level of disposable income, consumption equals 3.5 (in billions of dollars) plus 75 percent of disposable income.
 - a. What is the equation that expresses this relationship?
 - b. What is aggregate consumption when disposable income is 50 (in billions of dollars)?
 10. A firm has analyzed its sales and found that its customers will buy 20 percent more units of its products with each \$2 reduction in the unit price. When the price is \$12, the firm sells 500 units. What is the equation of the demand function for this product? Graph the equation.
 11.
 - a. Assume that water is in unlimited supply in a municipality. The consumer pays \$5.00 per month for water regardless of the amount used. Graph the demand and supply equations.
 - b. There is only one genuine painting by Rembrandt titled "The Night Watch." Assign arbitrary values to the painting and graph the demand and supply equations.
 12. A bus company has learned that when the price of a short excursion trip is \$5.00, 30 people will buy tickets; when the price is \$8.00, only 10 tickets will be sold. Obtain the point-slope form of the equation for the demand function and graph the equation.
 13. Identify which of the following equations represents a demand curve and which represents a supply curve; determine the equilibrium point and graph the curves.
 - a. $x + y = 5$
 - b. $2x - y = 5.5$
 14. Change equation b. in Problem 13 to $2x - y = 6$. Graph the equation and identify it as supply or demand. Has the equilibrium quantity involved increased or decreased relative to that in Problem 13?
 15. Suppose the fixed cost of production for a commodity is \$45,000. The variable cost is 60 percent of the selling price of \$15 per unit. What is the break-even quantity?
 16. If profit is \$100 per unit and the fixed cost of production is \$225,000, what is the break-even quantity?
 17. Aggregate national consumption is given (in billions of dollars) by the equation

$$c = 4.5 + 0.9y_d$$

where y_d is disposable income. If disposable income is 15 (in billions of dollars), a. What is aggregate consumption? b. What proportion of aggregate consumption is consumption of disposable income?

18. Aggregate national income (in billions of dollars) is 4.8 plus 80 percent of disposable income.
 - a. What is the equation of the consumption function?
 - b. What proportion of disposable income is consumed?
 - c. When disposable income is 60, what proportion of aggregate consumption is consumption of disposable income?

Answers to Odd-Numbered Problems

1. a. supply f. neither
 b. demand g. demand
 c. supply or demand h. supply
 d. supply or demand i. neither
 e. supply j. supply
3. a. i. 1, ii. 0.818, iii. 0.545 c. 0.091 (cannot be attained)
 b. i. 8.7, ii. 6.5, iii. 4.41
5. a. i. 5, ii. $\frac{a+3b}{a}$ c. b/a (cannot be attained)
 b. i. $2b$, ii. $4b$
7. a. \$25,000 d. break-even point at $x = 1000$;
 $TR = 5x$ recovery of fixed costs at
 c. \$13,000 $x = 600$
9. a. $c = 3.5 + 0.75y_d$ b. 41 billion dollars
13. equation a. is demand,
 equation b. is supply
 (3.5, 1.5)
15. 5000
17. a. 18 billion dollars b. 0.75

1.5 General Methods of Graphing Nonlinear Curves

Any equation whose graph is not a straight line is referred to as *nonlinear*. There are several types of nonlinear curves, all of which can be classified as either algebraic or transcendental.

A *polynomial* in x and y is the sum of a finite number of terms of the type $kx^r y^s$, where k is a constant and each of the exponents r and s is either a positive integer or zero. The *degree of a term* is $r + s$ and the *degree of the polynomial* is the degree of the term or terms of highest degree. Equations of the form $f(x, y) = 0$, where $f(x, y)$ is a polynomial in x and y , are referred to as *algebraic*; the graphical representation of an algebraic equation is an *algebraic curve*. Note that a straight line is an algebraic curve.

Any equation in x and y which is not algebraic is referred to as *transcendental*; the graphical representation of a transcendental equation is a *transcendental curve*. Included among the transcendental curves are the graphs of trigonometric, logarithmic, and exponential functions. Curves represented by algebraic equations of degree greater than two in x and y or by transcendental equations are referred to as *higher plane curves*.

The general methods of graphing discussed below can be used for graphing any type of algebraic or transcendental curve. However, for graphing many types of curves, special methods of graphing, which are more efficient than the general methods, are available. The general methods are discussed first. Then methods of identifying and graphing specific types of quadratic curves (circles, ellipses, parabolas, and hyperbolas) and trigonometric, exponential, and logarithmic curves are discussed.

Graphing a straight line from its equation is a relatively simple problem, since any two distinct points uniquely determine a straight line. However, graphing more

complicated equations is less straightforward. The graph representing any equation can be drawn fairly accurately if a sufficient number of points are plotted, but frequently so many points are required that this method is very laborious. Furthermore, plotting points in itself provides little or no information regarding the important properties of a curve.

From examining an equation, certain properties of the corresponding curve can be determined which will facilitate drawing the curve with a minimum of point plotting. Some of these properties are discussed below and their usefulness in graphing is illustrated. The properties discussed include intercepts, symmetry, extent, asymptotes, factorization, and real, point, and imaginary loci.

INTERCEPTS

The intercepts of a curve are the points at which it crosses the axes. The x -intercepts are obtained by setting $y = 0$ in the equation of the curve and solving for x ; the y -intercepts are obtained by setting $x = 0$ in the equation of the curve and solving for y .

SYMMETRY

Two points are *symmetric with respect to a line* if that line is the perpendicular bisector of the line segment joining the two points. Two points are symmetric with respect to a third point if that third point is the midpoint of the line segment joining the first two points. From these definitions it follows that the point (x, y) is a symmetric to the point

- $(x, -y)$ with respect to the x -axis
- $(-x, y)$ with respect to the y -axis
- $(-x, -y)$ with respect to the origin

as shown in Figure 1.29.

A curve may also be symmetric with respect to the x -axis, the y -axis or the origin, as shown in Figure 1.30. Refer to parts (a), (b), and (c) of that figure.

(a) A curve is symmetric to the x -axis if for each point (x, y) on the curve the symmetric point $(x, -y)$ is also on the curve—that is, if substitution of $-y$ for y does not change the form of the equation of the curve. For example, a curve is symmetric

Figure 1.29 Symmetry with respect to (a) the x -axis, (b) the y -axis, and (c) the origin.

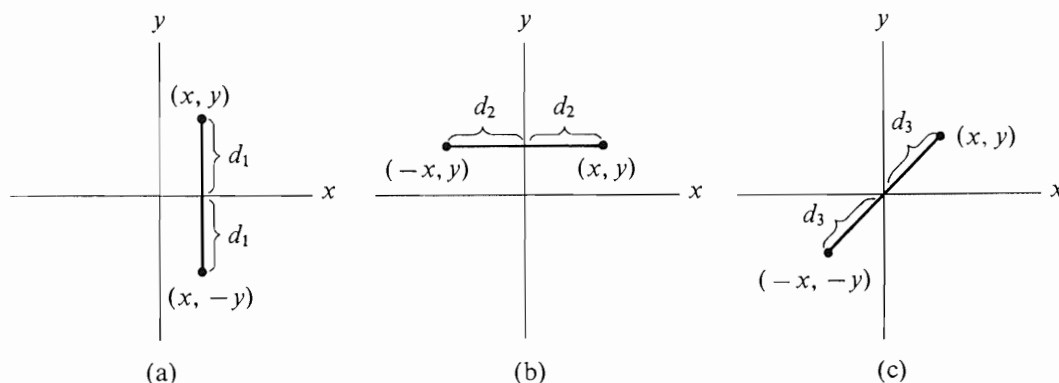
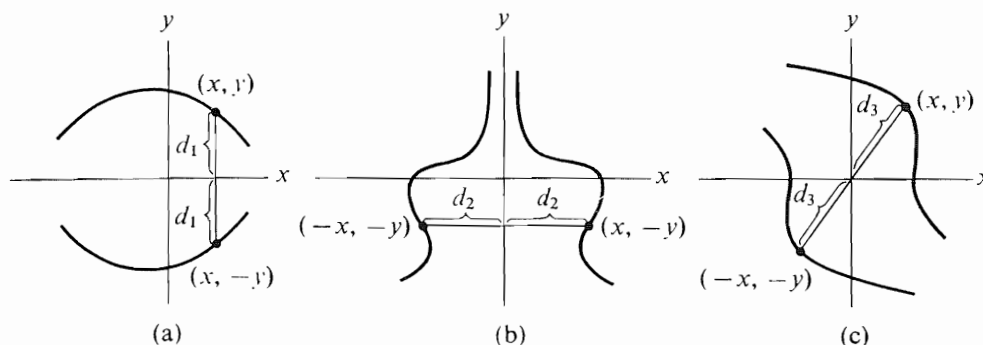


Figure 1.30 Symmetry with respect to (a) the x -axis, (b) the y -axis, and (c) the origin.

to the x -axis if it is represented by an algebraic equation in which y occurs only to even powers. Note, however, that there are algebraic curves symmetric to the x -axis whose equations contain odd powers of y . See the example which follows.

(b) A curve is symmetric to the y -axis if for each point (x, y) on the curve the symmetric point $(-x, y)$ is also on the curve—that is, if substitution of $-x$ for x does not change the form of the equation of the curve. For example, a curve is symmetric to the y -axis if it is represented by an algebraic equation in which x occurs only to even powers.

(c) A curve is symmetric to the origin if for each point (x, y) on the curve, the symmetric point $(-x, -y)$ is also on the curve—that is, if substitution of $-x$ for x and $-y$ for y does not change the form of the equation of the curve. For example, a curve is symmetric to the origin if it is represented by an algebraic equation each of whose terms is of even degree or each of whose terms is of odd degree.

In summary, the graph of an equation $f(x, y) = 0$ is symmetric with respect to

the x -axis if $f(x, y) = f(x, -y) = 0$

the y -axis if $f(x, y) = f(-x, y) = 0$

the origin if $f(x, y) = f(-x, -y) = 0$

Note that symmetry with respect to both the x - and y -axes implies symmetry about the origin, but symmetry about the origin does not imply symmetry about either axis. More generally, symmetry with respect to any two of the three (x -axis, y -axis, and origin) implies symmetry with respect to the third. Thus symmetry can occur with respect to none, one, or three of the x -axis, y -axis, and origin but never with respect to just two of them. This can be seen from the definitions above:

$$\left. \begin{array}{l} \text{Substitution of } -y \text{ for } y \\ \text{leaves the equation unchanged} \\ (\text{x-axis symmetry}) \\ \text{and} \\ \text{substitution of } -x \text{ for } x \\ \text{leaves the equation unchanged} \\ (\text{y-axis symmetry}) \end{array} \right\} \begin{array}{l} \Rightarrow \text{Substitution of } -x \text{ for } x \text{ and} \\ -y \text{ for } y \text{ leaves the equation} \\ \neq \text{unchanged (symmetry to origin)} \end{array}$$

and similarly for the other two cases. (The symbol \Rightarrow means “implies,” the symbol \neq means “does not imply,” the symbol \Leftarrow means “is implied by,” the symbol \nLeftarrow means “is not implied by.”)

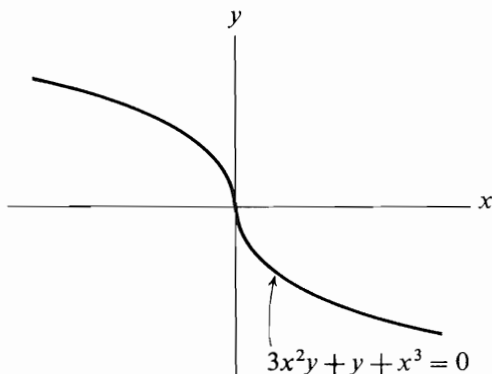
EXAMPLE

The curve represented by the equation

$$3x^2y + y + x^3 = 0$$

is symmetric to the origin but not to either axis (see Figure 1.31).

Figure 1.31



$f(x, -y) = -3x^2y - y + x^3$, so $f(x, -y) = 0$ is not the same equation as $f(x, y) = 0$ and $f(x, y) = 0$ is not symmetric to the x-axis.

$f(-x, y) = 3x^2y + y - x^3$, so $f(-x, y) = 0$ is not the same equation as $f(x, y) = 0$ and $f(x, y) = 0$ is not symmetric to the y-axis.

$f(-x, -y) = -3x^2y - y - x^3$, so $f(-x, -y) = 0$ is the same equation as $f(x, y) = 0$ and $f(x, y) = 0$ is symmetric to the origin.

EXAMPLE

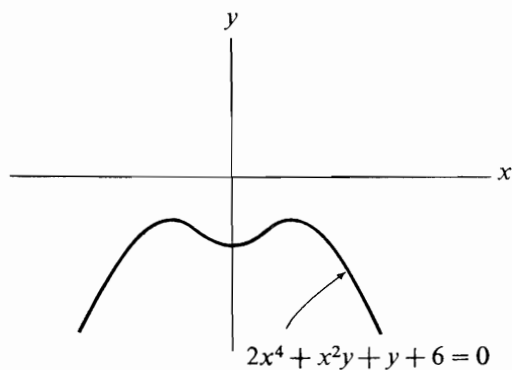
The curve represented by the equation

$$2x^4 + x^2y + y + 6 = 0$$

is symmetric to the y-axis but not to either the x-axis or the origin (see Figure 1.32).

$f(x, -y) = 2x^4 - x^2y - y + 6$, so $f(x, -y) = 0$ is not the same equation as $f(x, y) = 0$ and $f(x, y) = 0$ is not symmetric to the x-axis.

Figure 1.32



$f(-x, y) = 2x^4 + x^2y + y + 6$, so $f(-x, y) = 0$ is the same equation as $f(x, y) = 0$ and $f(x, y) = 0$ is symmetric to the y -axis.

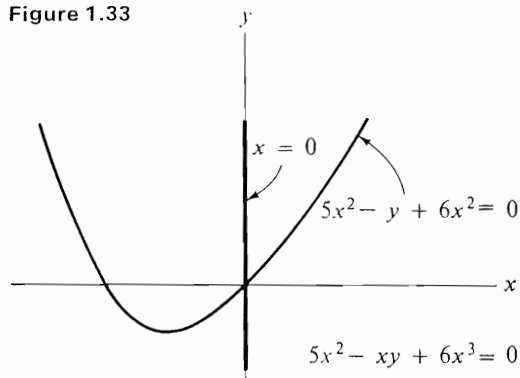
$f(-x, -y) = 2x^4 - x^2y - y + 6$, so $f(-x, -y) = 0$ is not the same equation as $f(x, y) = 0$ and $f(x, y) = 0$ is not symmetric to the origin.

EXAMPLE

The curve represented by the equation

$$5x^2 - xy + 6x^3 = 0$$

is not symmetric to the x -axis, the y -axis, or the origin (see Figure 1.33).

Figure 1.33

$f(x, -y) = 5x^2 + xy + 6x^3$, so $f(x, -y) = 0$ is not the same equation as $f(x, y) = 0$ and $f(x, y) = 0$ is not symmetric to the x -axis.

$f(-x, y) = 5x^2 + xy - 6x^3$, so $f(-x, y) = 0$ is not the same equation as $f(x, y) = 0$ and $f(x, y) = 0$ is not symmetric to the y -axis.

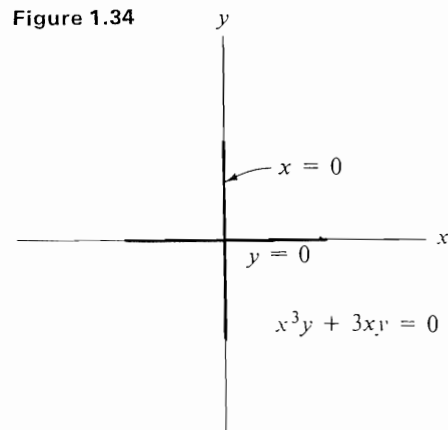
$f(-x, -y) = 5x^2 - xy - 6x^3$, so $f(-x, -y) = 0$ is not the same equation as $f(x, y) = 0$ and $f(x, y) = 0$ is not symmetric to the origin.

EXAMPLE

The curve represented by the equation

$$x^3y + 3xy = 0$$

is symmetric to the x -axis, the y -axis, and the origin (see Figure 1.34).

Figure 1.34

$f(x, -y) = -x^3y - 3xy$, so $f(x, -y) = 0$ is the same equation as $f(x, y) = 0$ and $f(x, y) = 0$ is symmetric to the x -axis.

$f(-x, y) = -x^3y - 3xy$, so $f(-x, y) = 0$ is the same equation as $f(x, y) = 0$ and $f(x, y) = 0$ is symmetric to the y -axis.

$f(-x, -y) = x^3y + 3xy$, so $f(-x, -y) = 0$ is the same equation as $f(x, y) = 0$ and $f(x, y) = 0$ is symmetric to the origin.

In graphing an equation it is seldom practical to consider symmetry with respect to other than the coordinate axes and the origin; exceptions to this are curves whose equations are of a particular form from which lines or points of symmetry are readily identified. Examples of such equations are given in later sections.

Occasionally it is desired to construct an equation whose curve is symmetric with respect to certain lines or points. In these problems the following rules, of which the above are special cases, are useful (see Figure 1.35): The graph of the equation $f(x, y) = 0$ is symmetric with respect to

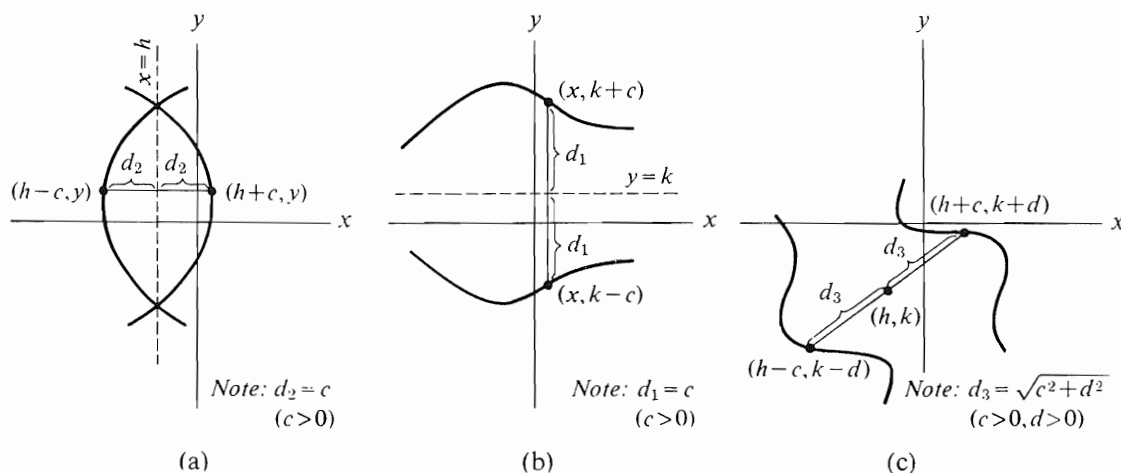
the line $x = h$ if $f(h + c, y) = f(h - c, y) = 0$ for all c and y

the line $y = k$ if $f(x, k + c) = f(x, k - c) = 0$ for all c and x

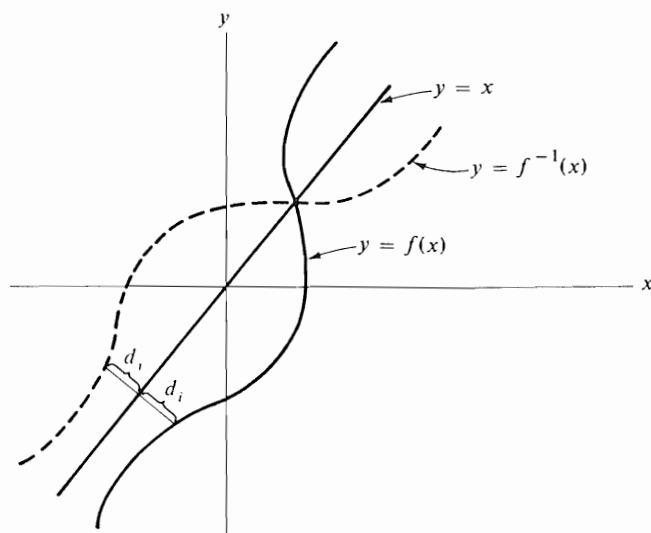
the point (h, k) if $f(h + c, k + d) = f(h - c, k - d) = 0$ for all c and d

NOTE: Symmetry about the line $y = mx + b$ must be defined in terms of transformations. This procedure is discussed in later sections.

Figure 1.35 Symmetry with respect to (a) the line $x = h$, (b) the line $y = k$, and (c) the point (h, k) .



Symmetry with respect to the line $y = x$ is a property of inverse functions. Geometrically, the problem of finding the inverse function of $f(x)$ is the problem of viewing the graph of $y = f(x)$ with the x -axis and y -axis interchanged, which is equivalent to reflection about the line $y = x$. Thus the graph of the inverse of a function can be obtained by drawing a perpendicular from each point on the graph of the original function to the line $y = x$ and extending the perpendicular an equal distance beyond the line $y = x$. The points thus obtained are the graph of the inverse function $y = f^{-1}(x)$. (See Figure 1.36.)

Figure 1.36 Symmetry with respect to the line $y = x$.**EXTENT**

The rectangular coordinate system defined represents only those points (x, y) whose coordinates are real numbers; thus values of x for which y is imaginary and values of y for which x is imaginary must be excluded. When even powers of a variable appear in an equation, the solution for that variable may involve square roots (or other even roots) and the extent of the curve may then be limited, since negative numbers do not have real square roots (or other even roots). When possible, an equation should be solved explicitly for each variable to see whether the range of any variable is limited and the curve is correspondingly limited in extent. Note that any limitation on the extent of one variable may implicitly limit the extent of the other variable.

EXAMPLE

Determine whether there is any limitation on the extent of the curve represented by the equation

$$x^2 + y^2 = 9$$

Solving for x ,

$$x = \pm\sqrt{9 - y^2}$$

the quantity under the radical sign $9 - y^2$ is negative if $|y| > 3$; the extent of the curve in the y -direction is therefore limited to the interval $-3 \leq y \leq 3$.

Solving for y ,

$$y = \pm\sqrt{9 - x^2}$$

and the extent of the curve in the x -direction is therefore similarly limited to the interval $-3 \leq x \leq 3$.

EXAMPLE

Determine whether there is any limitation on the extent of the curve represented by the equation

$$x^2 - y^2 = 4$$

Solving for x ,

$$x = \pm \sqrt{4 + y^2}$$

the quantity under the radical sign $4 + y^2$ is always positive and the extent of the curve in the y -direction is therefore not limited.

Solving for y ,

$$y = \pm \sqrt{x^2 - 4}$$

the quantity under the radical sign $x^2 - 4$ is negative if $|x| < 2$; the extent of the curve in the x -direction is therefore limited to the values $-2 \leq x \leq 2$.

PROBLEMS

For each of the following equations, determine a. the intercepts; b. whether the curve represented is symmetric to the x -axis, the y -axis, or the origin; and c. whether there is any limitation on extent.

1. $x^3 - y^2 - 9 = 0$

2. $x^2 + y^2 - 18 = 0$

3. $y^2 - 2x + 5 = 0$

4. $xy + 5x - 15 = 0$

5. $x^2 + y^4 - 6 = 0$

6. $x^2y^2 - 25 = 0$

Answers to Odd-Numbered Problems

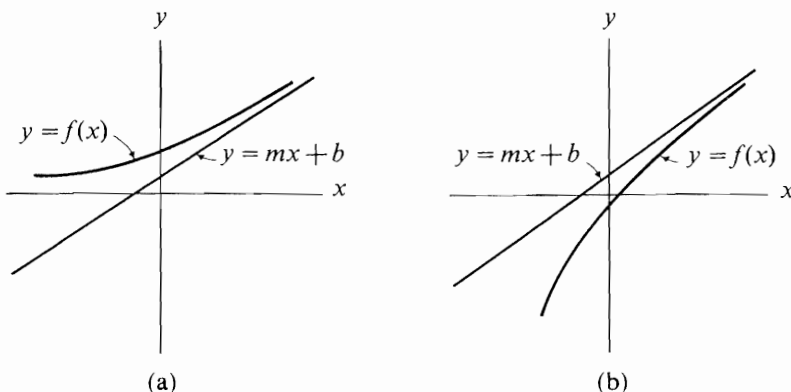
1. a. intercepts: $(\sqrt[3]{9}, 0)$
 b. symmetry: about the x -axis
 c. extent: $x^3 \geq 9$, y not limited
3. a. intercepts: $(\frac{5}{2}, 0)$
 b. symmetry: about the x -axis
 c. extent: $x \geq \frac{5}{2}$, y not limited
5. a. intercepts: $(0, \sqrt[4]{6})$, $(\sqrt[4]{6}, 0)$
 b. symmetry: about the x -axis, y -axis, and origin
 c. extent: $x^2 \leq 6$, $y^4 \leq 6$

ASYMPTOTES

A curve may have the property that a point can move along it so that the distance from the origin to the point increases without limit. If, as the distance (in some direction) of a point from the origin increases without limit, the distance between the point and a fixed straight line approaches zero through values of the same sign, the line is called an *asymptote* of the curve.

To show that a particular line is an asymptote of a particular curve it is necessary to show that the curve approaches arbitrarily close to the line as the distance from the origin increases without bound. This property involves the concept of limits, which will be discussed in succeeding sections.

In general, then, the line $y = mx + b$ is an asymptote of the curve $y = f(x)$ if $f(x)$ becomes arbitrarily close to $mx + b$ as x and y increase without bound. This may be written $f(x) \rightarrow mx + b$ as $x, y \rightarrow \infty$. (See Figure 1.37.)

Figure 1.37 $f(x) \rightarrow mx + b$ as $x, y \rightarrow \infty$; (a) from above, (b) from below.

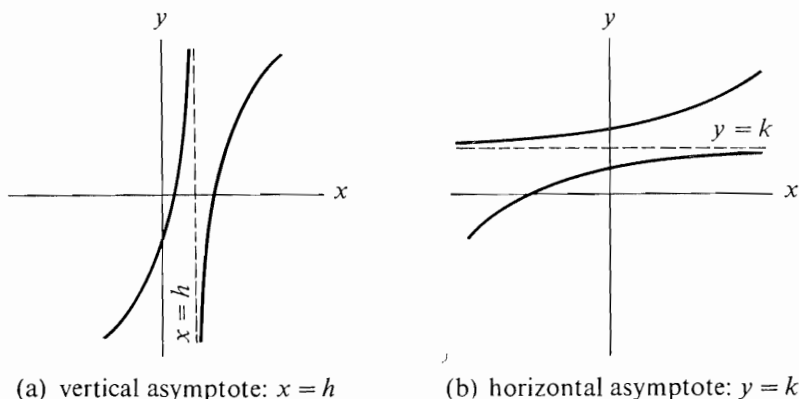
Asymptotes that are parallel to or coincident with one of the coordinate axes are frequently of particular interest. These vertical and horizontal asymptotes are defined as follows (see Figure 1.38).

The line $x = h$ is a *vertical asymptote* of the curve $y = f(x)$ if $x \rightarrow h$ as $y \rightarrow \infty$.

The line $y = k$ is a *horizontal asymptote* of the curve $y = f(x)$ if $y \rightarrow k$ as $x \rightarrow \infty$.

For purposes of graphing, it is helpful to determine the behavior of a curve with respect to each of its asymptotes, in addition to determining the existence of the asymptotes. Thus the equation should be examined as x and y in turn increase without bound ($x \rightarrow +\infty, y \rightarrow +\infty$) and as x and y in turn decrease without bound ($x \rightarrow -\infty, y \rightarrow -\infty$); the value of the variable which is not increasing or decreasing without bound should also be noted, in order to determine whether the curve is approaching its asymptote from the left or the right (vertical asymptotes) or from above or below (horizontal asymptotes). When investigating asymptotes, it is usually helpful to solve an equation explicitly first for one variable and then for the other variable, if possible. As noted above, this procedure is also useful in determining possible limitations on extent; since such limitations prevent the variable involved from increasing or decreasing without bound, they are important in considering asymptotes.

Figure 1.38



EXAMPLE

Determine whether the curve represented by the following equation has vertical or horizontal asymptotes:

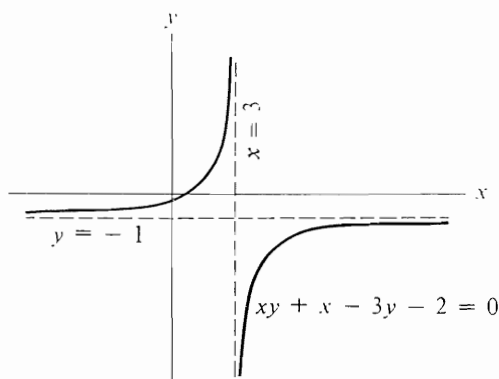
$$xy + x - 3y - 2 = 0$$

(see Figure 1.39).

Solving for x ,

$$x = \frac{3y + 2}{y + 1}$$

Figure 1.39



as $y \rightarrow +\infty$, $x \rightarrow 3$, and $x < 3$; as $y \rightarrow -\infty$, $x \rightarrow 3$, and $x > 3$; thus $x = 3$ is an asymptote of the curve and is approached from left and right.

Solving for y ,

$$y = \frac{x - 2}{3 - x}$$

as $x \rightarrow +\infty$, $y \rightarrow -1$, and $y < -1$; as $x \rightarrow -\infty$, $y \rightarrow -1$, and $y > -1$; thus $y = -1$ is an asymptote of the curve and is approached from above and below.

EXAMPLE

Determine whether the curve represented by the following equation has vertical or horizontal asymptotes:

$$x^2 - y - 1 = 0$$

(see Figure 1.40).

Solving for x ,

$$x = \pm\sqrt{y + 1}$$

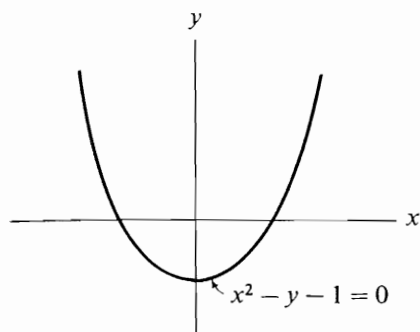
as $y \rightarrow +\infty$, $x \rightarrow \pm\infty$; $y > -1$, so y cannot approach $-\infty$; thus the curve does not have a vertical asymptote.

Solving for y ,

$$y = x^2 - 1$$

as $x \rightarrow +\infty$, $y \rightarrow +\infty$; as $x \rightarrow -\infty$, $y \rightarrow +\infty$ (as above); thus the curve does not have a horizontal asymptote.

Figure 1.40

**FACTORIZATION**

If the left member of the equation $f(x, y) = 0$ can be written as the product of two or more factors, for example, if $f(x, y) = g(x, y) \cdot h(x, y) = 0$, then only those points (x, y) whose coordinates satisfy either $g(x, y) = 0$ or $h(x, y) = 0$ lie on the graph of $f(x, y) = 0$. Thus the graph of $f(x, y) = 0$ consists of the graphs of $g(x, y) = 0$ and $h(x, y) = 0$. It is important to check for factorization, since a factorable equation may be difficult to graph correctly if it is not factored.

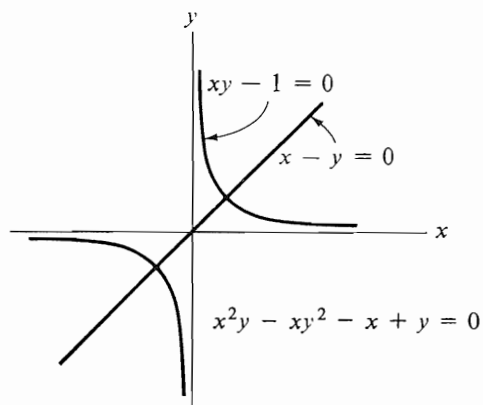
EXAMPLE

Graph the equation

$$x^2y - xy^2 - x + y = 0$$

(see Figure 1.41).

Figure 1.41



Factoring,

$$xy(x - y) - x + y = 0$$

$$xy(x - y) - (x - y) = 0$$

$$(xy - 1)(x - y) = 0$$

Thus the graph of $x^2y - xy^2 - x + y = 0$ consists of the graphs of the curve $xy - 1 = 0$ and the straight line $x - y = 0$.

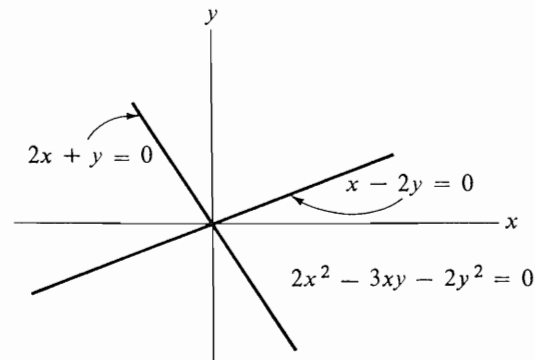
EXAMPLE

Graph the equation

$$2x^2 - 3xy - 2y^2 = 0$$

(see Figure 1.42).

Figure 1.42



Factoring,

$$(2x + y)(x - 2y) = 0$$

Thus the graph of $2x^2 - 3xy - 2y^2 = 0$ consists of the graphs of the straight lines $2x + y = 0$ and $x - 2y = 0$.

REAL CURVE, POINT LOCUS, OR IMAGINARY LOCUS

Some equations are satisfied by the coordinates of only one point or of a finite number of points; the graphs of such equations are called *point loci*. Other equations are satisfied by the coordinates of no real point; since such equations have no graphical representation in a system of real axes, they are said to represent *imaginary loci*.

EXAMPLE

The equation $x^2 + y^2 = 0$ is satisfied only by the point $(0, 0)$, and its graph is thus a point locus.

EXAMPLE

The equation $(x^2 - 4)^2 + (y^2 - 4)^2 = 0$ is satisfied only by the points $(2, 2)$, $(2, -2)$, $(-2, 2)$, and $(-2, -2)$, and its graph is thus a point locus.

EXAMPLE

The equation $x^2 + y^2 = -5$ is satisfied by no pair of real numbers x and y and its locus is thus imaginary.

PROBLEMS

For each of the following equations, determine a. the asymptotes; b. whether the equation can be factored; and c. whether the equation represents a real curve, point locus, or imaginary curve.

- | | |
|---------------------------|-------------------------------|
| 1. $2xy - x + y + 5 = 0$ | 4. $3x^2 + 3xy - 2x - 2y = 0$ |
| 2. $3x^2 + 2xy - y^2 = 0$ | 5. $3x^2 + 6y^2 + x^2y^2 = 0$ |
| 3. $2x^2 + 3y^2 + 6 = 0$ | 6. $3x^2 - 4y^2 - 9 = 0$ |

Answers to Odd-Numbered Problems

1. a. asymptotes: $y \rightarrow \infty$, $x \rightarrow -\frac{1}{2}$; $x \rightarrow \infty$, $y \rightarrow \frac{1}{2}$
 b. cannot be factored
 c. real locus
3. a. asymptotes: none
 b. cannot be factored
 c. imaginary locus
5. a. asymptotes: none
 b. cannot be factored
 c. point locus

SUMMARY: GENERAL METHODS OF GRAPHING NONLINEAR CURVES

In the preceding sections properties of curves that are helpful in graphing are discussed. These properties include:

- | | |
|---------------|--|
| 1. Intercepts | 4. Asymptotes |
| 2. Symmetry | 5. Factorization |
| 3. Extent | 6. Real curve, point locus, or imaginary locus |

The following examples illustrate the use of these six properties in graphing a given equation. Although factorization occurs near the end of the list (because in practice factorable equations occur relatively infrequently), it should be checked first, as it may make investigation of some of the other properties unnecessary. The

emphasis here is on determining properties of the curves from their equations; greater accuracy in graphing can be attained by plotting more points and, particularly, by methods of differential calculus to be discussed subsequently.

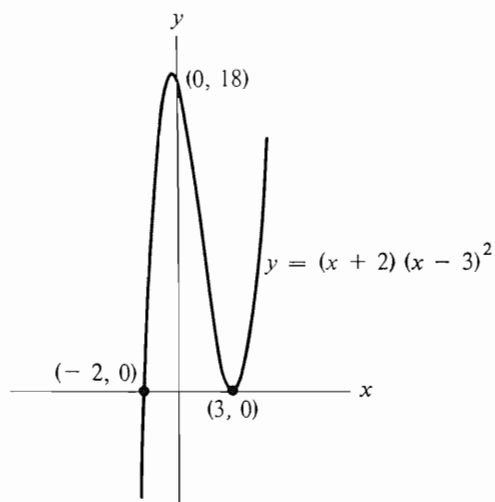
EXAMPLE

Graph the equation

$$y = (x + 2)(x - 3)^2$$

(see Figure 1.43).

Figure 1.43



Intercepts: If $y = 0$, $x = -2, 3$; if $x = 0$, $y = 18$.

Symmetry: Not symmetric with respect to either axis or the origin.

Extent: No apparent limits.

Asymptotes: None (as $x \rightarrow +\infty$, $y \rightarrow +\infty$ and as $x \rightarrow -\infty$, $y \rightarrow -\infty$).

Factorization: Not factorable. [Note that factorability refers to the equation $f(x, y) = (x + 2)(x - 3)^2 - y = 0$, which is not factorable, although $y = f(x) = (x + 2)(x - 3)^2$ is, in fact, factored.]

Real, point, or imaginary locus: Real curve.

EXAMPLE

Graph the equation

$$x^2 - 4x + y - 12 = 0$$

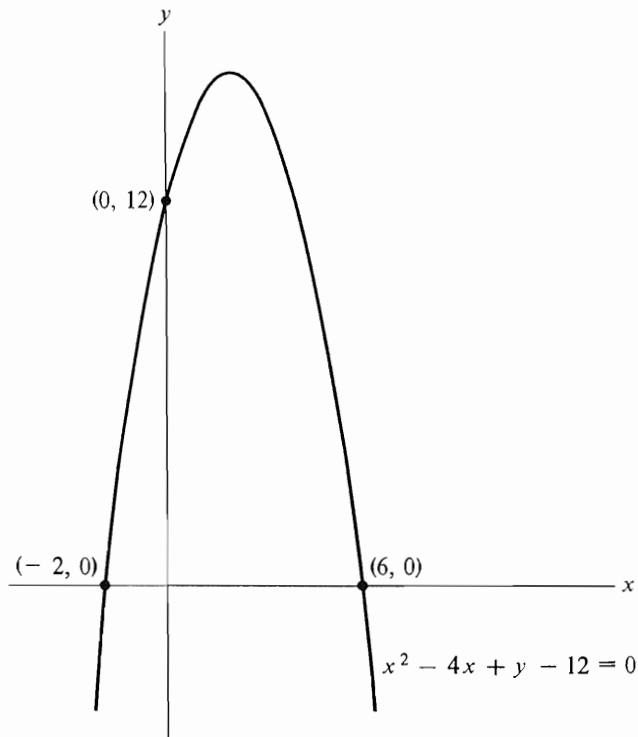
(see Figure 1.44).

Intercepts: If $y = 0$, $x = -2, 6$; if $x = 0$, $y = 12$.

Symmetry: Not symmetric with respect to either axis or the origin.

Extent: No apparent limits.

Figure 1.44



Asymptotes: None (as $x \rightarrow +\infty$, $y \rightarrow -\infty$ and $x \rightarrow -\infty$, $y \rightarrow -\infty$).

Factorization: Not factorable.

Real, point, or imaginary locus: Real curve.

EXAMPLE

Graph the equation

$$xy - y - x - 2 = 0$$

(see Figure 1.45).

Intercepts: If $y = 0$, $x = -2$; if $x = 0$, $y = -2$.

Symmetry: Not symmetric with respect to either axis or the origin.

Extent: Solving for x ,

$$x = \frac{y + 2}{y - 1}$$

So no limitation on the extent of y , except x is not defined for $y = 1$. Solving for y ,

$$y = \frac{x + 2}{x - 1}$$

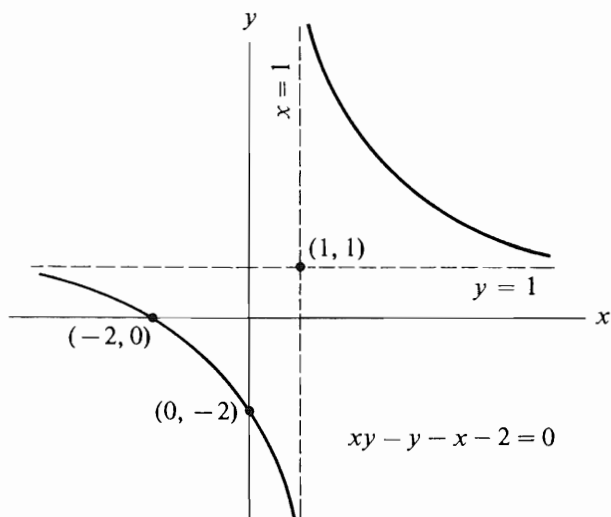
So no limitation on the extent of x except y is not defined for $x = 1$.

Asymptotes: $x = (y + 2)/(y - 1)$. As $y \rightarrow +\infty$, $x \rightarrow 1$ and $x > 1$; as $y \rightarrow -\infty$, $x \rightarrow 1$ and $x < 1$; so $x = 1$ is an asymptote. $y = (x + 2)/(x - 1)$. As $x \rightarrow +\infty$, $y \rightarrow 1$ and $y > 1$; as $x \rightarrow -\infty$, $y \rightarrow 1$ and $y < 1$; so $y = 1$ is an asymptote.

Factorization: Not factorable.

Real, point, or imaginary locus: Real curve.

Figure 1.45

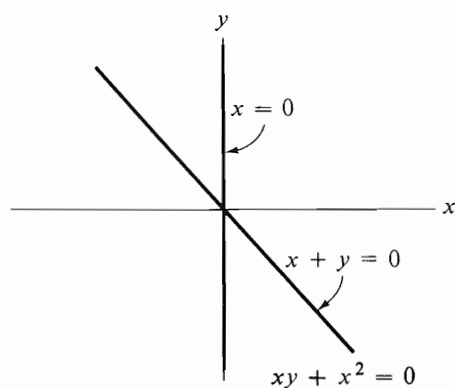
**EXAMPLE**

Graph the equation

$$xy + x^2 = 0$$

(see Figure 1.46).

Figure 1.46



$xy + x^2 = x(y + x) = 0$, so the graph consists of the two straight lines $x = 0$ and $x + y = 0$. The line $x = 0$ is the y -axis. The line $x + y = 0$ has intercept $(0, 0)$ and slope -1 .

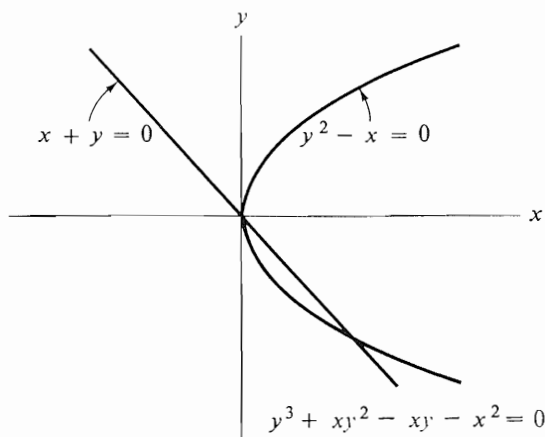
EXAMPLE

Graph the equation

$$y^3 + xy^2 - xy - x^2 = 0$$

(see Figure 1.47).

Figure 1.47



$y^3 + xy^2 - xy - x^2 = y^2(y + x) - x(y + x) = (y^2 - x)(y + x) = 0$, so the graph consists of the straight line $x + y = 0$ and the curve $y^2 - x = 0$. The line $x + y = 0$ has intercept $(0, 0)$ and slope -1 . Graphing the equation $y^2 - x = 0$,

Intercepts: If $y = 0$, $x = 0$; if $x = 0$, $y = 0$.

Symmetry: Symmetric with respect to the x -axis.

Extent: Solving for x , $x = y^2$, so there is no limitation on the extent of y ; solving for y , $y = \pm\sqrt{x}$, so x is limited to $x \geq 0$.

Asymptotes: None (as $x \rightarrow +\infty$, $y \rightarrow \pm\infty$).

EXAMPLE

Graph the equation

$$x^4 + x^2 - y^2 - y = 0$$

(see Figure 1.48).

$$\begin{aligned} x^4 + x^2 - y^2 - y &= (x^2 + y)(x^2 - y) + (x^2 - y) \\ &= (x^2 - y)(x^2 + y + 1) = 0 \end{aligned}$$

so the graph consists of the curves $x^2 - y = 0$ and $x^2 + y + 1 = 0$. Graphing the equation $x^2 - y = 0$,

Intercepts: If $y = 0$, $x = 0$; if $x = 0$, $y = 0$.

Symmetry: Symmetric with respect to the y -axis.

Extent: Solving for x , $x = \pm\sqrt{y}$, so y is limited to $y \geq 0$. Solving for y , $y = x^2$, so there is no limitation on the extent of x .

Asymptotes: None (as $x \rightarrow +\infty$, $y \rightarrow +\infty$; as $x \rightarrow -\infty$, $y \rightarrow +\infty$).

Graphing the equation $x^2 + y + 1 = 0$,

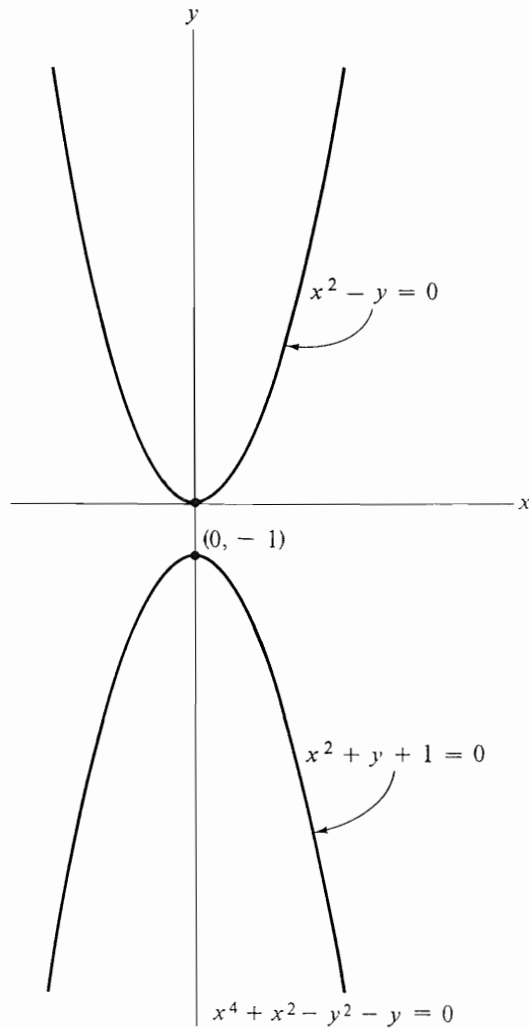
Intercepts: If $y = 0$, $x = \sqrt{-1}$ (imaginary); if $x = 0$, $y = -1$.

Symmetry: Symmetric with respect to the y -axis.

Extent: Solving for x , $x = \pm\sqrt{-y-1}$, so y is limited to $y \leq -1$. Solving for y , $y = -1 - x^2$, so there is no limitation on the extent of x .

Asymptotes: None (as $x \rightarrow +\infty$, $y \rightarrow +\infty$; as $x \rightarrow -\infty$, $y \rightarrow +\infty$).

Figure 1.48



EXAMPLE

Graph the equation

$$x^2y - x^2 - 4y = 0$$

(see Figure 1.49).

Intercepts: If $y = 0$, $x = 0$; if $x = 0$, $y = 0$.

Symmetry: Symmetric with respect to the y -axis.

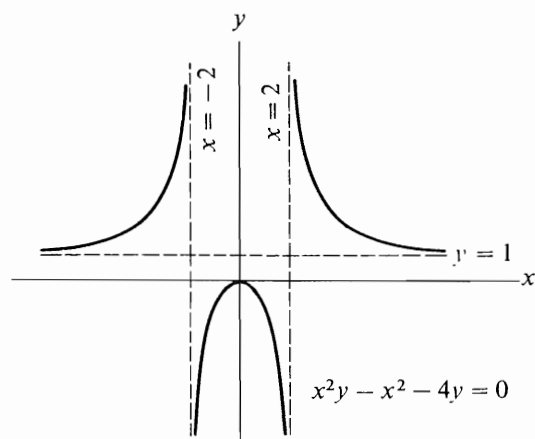
Extent: Solving for x ,

$$x = \pm \sqrt{\frac{4y}{y-1}} = \pm 2\sqrt{\frac{y}{y-1}}$$

$y/(y-1)$ is negative if $0 < y < 1$ and is undefined if $y = 1$, so y is limited to the intervals $y \leq 0$ and $y > 1$. Solving for y , $y = x^2/(x^2 - 4)$, so there is no limitation on the extent of x except y is not defined for $x = \pm 2$.

Asymptotes: $x = \pm 2\sqrt{y/(y-1)}$. As $y \rightarrow +\infty$, $x \rightarrow \pm 2$ and $x > |2|$; as $y \rightarrow -\infty$, $x \rightarrow \pm 2$

Figure 1.49



and $x < |2|$; so $x = \pm 2$ are asymptotes. $y = x^2/(x^2 - 4)$. As $x \rightarrow +\infty$, $y \rightarrow 1$ and $y > 1$; as $x \rightarrow -\infty$, $y \rightarrow 1$ and $y > 1$, so $y = 1$ is an asymptote.

Factorization: Not factorable.

Real, point, or imaginary locus: Real curve.

EXAMPLE

Graph the equation

$$x^2y - x^2 + 4y = 0$$

(see Figure 1.50).

Intercepts: If $x = 0$, $y = 0$; if $y = 0$, $x = 0$.

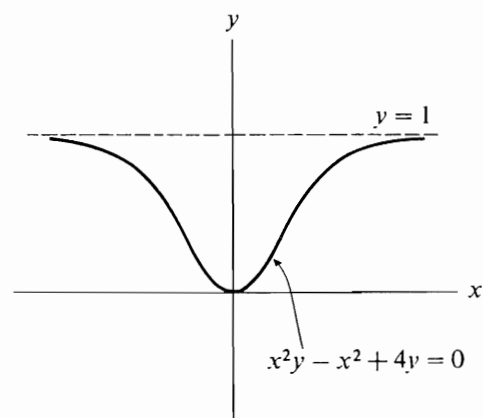
Symmetry: Symmetric with respect to the y -axis.

Extent: Solving for x ,

$$x = \pm \sqrt{\frac{4y}{1-y}} = \pm 2\sqrt{\frac{y}{1-y}}$$

$y/(1-y)$ is negative if $y < 0$ or $y > 1$ and is undefined if $y = 1$, so y is limited to the interval $0 \leq y < 1$. Solving for y , $y = x^2/(x^2 + 4)$, so there is no limitation on the extent of x .

Figure 1.50



Asymptotes: $x = \pm 2\sqrt{y/(1-y)}$. y is limited in extent and cannot become large or small without bound. $y = x^2/(x^2 + 4)$. As $x \rightarrow +\infty$, $y \rightarrow 1$ and $y < 1$; as $x \rightarrow -\infty$, $y \rightarrow 1$ and $y < 1$, so $y = 1$ is an asymptote.

Factorization: Not factorable.

Real, point, or imaginary locus: Real curve.

EXAMPLE

Graph the equation

$$y^2 - 4xy^2 - 1 = 0$$

(see Figure 1.51).

Intercepts: If $y = 0$, x is not defined; if $x = 0$, $y = \pm 1$.

Symmetry: Symmetric with respect to the x -axis.

Extent: Solving for x ,

$$x = \frac{y^2 - 1}{4y^2}$$

so there is no limitation on the extent of y , except that x is not defined for $y = 0$. Solving for y ,

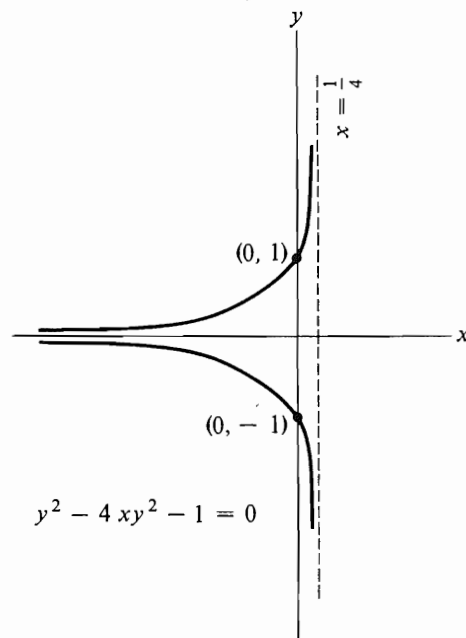
$$y = \pm \sqrt{\frac{1}{1-4x}}$$

so x is limited to $x < \frac{1}{4}$.

Asymptotes: $x = (y^2 - 1)/4y^2$. As $y \rightarrow \pm\infty$, $x \rightarrow \frac{1}{4}$, so $x = \frac{1}{4}$ is an asymptote.

$$y = \pm \sqrt{\frac{1}{1-4x}}$$

Figure 1.51



x is limited in extent and cannot become large without bound. As $x \rightarrow -\infty$, $y \rightarrow 0$, so $y = 0$ is an asymptote.

Factorization: Not factorable.

Real, point, or imaginary locus: Real curve.

PROBLEMS

Sketch the curves represented by the following equations; give intercepts, extent, symmetry, and asymptotes, as relevant.

- | | |
|-----------------------------|-----------------------------------|
| 1. $x^2y = 10$ | 8. $y = x^3(x - 1)(x + 6)$ |
| 2. $xy^2 = -10$ | 9. $4y = x^3$ |
| 3. $y = x(x - 3)(x + 4)$ | 10. $y = x^2(x - 3)^2$ |
| 4. $y = x^2(x^2 - 4x + 4)$ | 11. $y = x\sqrt{9 - x^2}$ |
| 5. $y = x^4 - x^2$ | 12. $y = x^2\sqrt{9 - x^2}$ |
| 6. $y = (x^2 - 1)(x^2 - 4)$ | 13. $y = (x - 3)(x^2 - 4x - 5)$ |
| 7. $y = x^3 - 4x$ | 14. $y = x^2(x - 6)(x^2 - x - 6)$ |

Answers to Odd-Numbered Problems

- | | |
|--|---|
| 1. intercepts: none
extent: $y > 0$, x not limited
symmetry: about y -axis
asymptotes: $x = 0$, $y = 0$ | 9. intercepts: $(0, 0)$
extent: not limited
symmetry: about origin
asymptotes: none |
| 3. intercepts: $(0, 0)$, $(3, 0)$, $(-4, 0)$
extent: not limited
symmetry: none
asymptotes: none | 11. intercepts: $(0, 0)$, $(\pm 3, 0)$
extent: $-3 \leq x \leq 3$, y not limited
except by range of x
symmetry: about origin
asymptotes: none |
| 5. intercepts: $(0, 0)$, $(\pm 1, 0)$
extent: not limited
symmetry: about y -axis
asymptotes: none | 13. intercepts: $(0, 15)$, $(3, 0)$, $(5, 0)$,
$(-1, 0)$
extent: not limited
symmetry: none
asymptotes: none |
| 7. intercepts: $(0, 0)$, $(\pm 2, 0)$
extent: not limited
symmetry: about origin
asymptotes: none | |

1.6 Quadratic Curves

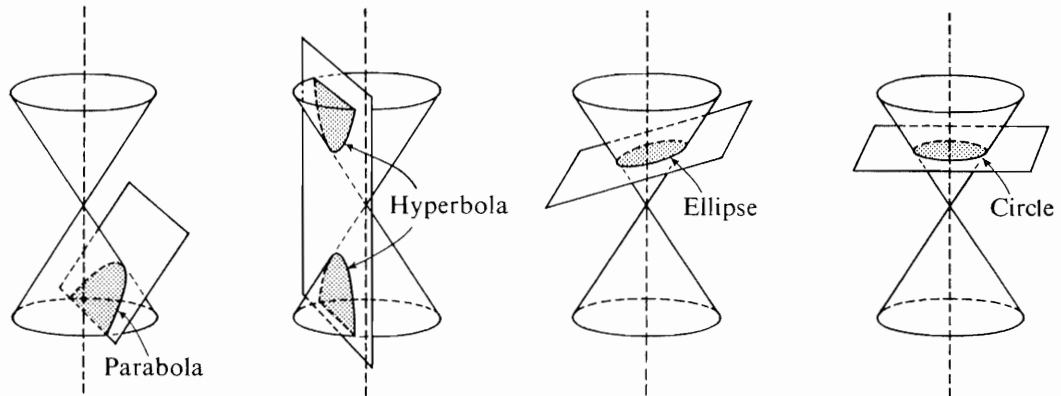
A quadratic equation may represent a circle, an ellipse, a parabola, a hyperbola, or various degenerate or imaginary cases of these curves. The identification of a quadratic equation as representing one of these types of curves is discussed in this section; the standard form for each type of equation is then considered for purposes of graphical representation. Families of quadratic curves are also discussed.

An equation of the form

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0 \quad (2)$$

where A , B , C , D , E , and F are constants and at least one of A , B , and C is nonzero, is said to be a *second-degree* or *quadratic equation*. An equation is said to be quadratic in a variable if that variable occurs to the second and no higher power. Equation (2) is the general equation of second degree in two variables.

Figure 1.52 Conic sections.



Since it is possible to remove the xy -term from any second-degree equation, making B zero, the general quadratic equation actually can be written

$$Ax^2 + Cy^2 + Dx + Ey + F = 0$$

The removal of the xy -term is accomplished by changing from one set of axes to a new set, which can be obtained by rotating the original axes about the origin. This procedure will not be discussed in detail, because it is seldom necessary to use it in practice.

Essentially, absence of the xy -term in a second-degree equation indicates that the corresponding curve is symmetric to a line or lines (axes of the curve) parallel to one or both of the coordinate axes. A circle, by definition, always has this type of symmetry and B is always zero for a circle.

The curves represented by Equation (2) can be obtained by cutting a (double) cone with a plane and thus are known as *conic sections*. The four special forms of a conic section are the circle, the ellipse, the parabola, and the hyperbola (see Figure 1.52). In addition to these curves, Equation (2) may represent various types of *degenerate conics* (actually special cases of conic sections). Degenerate conics may be represented by two intersecting straight lines, two parallel straight lines, two coincident straight lines (that is, one straight line), a point, or no points at all.

Any conic section can be defined as the locus of a point that moves in such a way that the ratio of its distance from a fixed point (called its *focus*) to its distance from a fixed line (called its *directrix*) is a constant e (called its *eccentricity*). Conic sections can be classified into the four basic types according to the value of their eccentricity. (For a circle, $e = 0$; for an ellipse, $0 < e < 1$; for a parabola, $e = 1$; for a hyperbola, $e > 1$.)

There are many interesting scientific applications of conic sections. For example, the planets move in elliptic orbits about the sun with the sun at a focus; whispering galleries usually have elliptical ceilings arranged so that a person standing at one focus can hear a slight noise made at the other focus while someone standing between the two foci hears nothing; elliptic arches are frequently used in architecture. The cable of a suspension bridge hangs in the form of an arc of a parabola; a reflecting surface made by rotating a parabola about its axis will send the light out in parallel rays if the source of light is at the focus; the larger reflector in a reflecting telescope is usually parabolic; the path of a projectile, neglecting air resistance, is a parabola. The small reflector of a reflecting telescope is often a hyperbolic mirror (used to reflect the image to the eyepiece); the hyperbola can be used to locate an invisible source of

sound, such as an enemy's guns, by considering two listening posts as the foci of a hyperbola and performing certain computations on the difference in time of arrival of the sound of the gun at the listening posts.

In economics and business, certain types of parabolic and hyperbolic curves are appropriate for representing demand and supply functions, production functions, and many other relationships. Applications of circular and elliptic curves are less common but occur, for example, in representing product transformation functions.

The type of conic section represented by a second-degree equation can be identified from an examination of its coefficients, whether or not the equation includes an xy -term. However, only if the equation has no xy -term can it readily be put into an appropriate standard form from which properties useful for graphing and for other purposes can be obtained and from which degenerate cases can be identified.

Identification of a Quadratic Equation

The general quadratic equation

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$$

with at least one of A and C nonzero can be identified as a circle, ellipse, parabola, or hyperbola as follows:

- If $B = 0$ and $A = C$, the conic is a circle
- If $B^2 - 4AC < 0$, the conic is an ellipse
- If $B^2 - 4AC = 0$, the conic is a parabola
- If $B^2 - 4AC > 0$, the conic is a hyperbola

Direct identification of degenerate and imaginary loci is discussed in Technical Note II.

For the special case $B = 0$, with at least one of A and C nonzero, the above identification procedure can be reduced to the following:

- If $A = C$, the conic is a circle
- If $A \neq C$, but A and C have the same sign, the conic is an ellipse
- If $A = 0$ or $C = 0$ but not both, the conic is a parabola
- If A and C have opposite signs, the conic is a hyperbola

The standard forms for each of the four types of conic sections and their geometrical representations and properties are discussed in the following sections. The algebraic procedure known as "completing the square," which is used to put equations into these standard forms, is discussed in detail in Technical Note III.

THE CIRCLE

Geometrically, a circle is the locus of points in a plane which are a fixed distance from a given point called the *center*. The distance of the points from the center is the *radius* of the circle. The general equation of a circle can be written

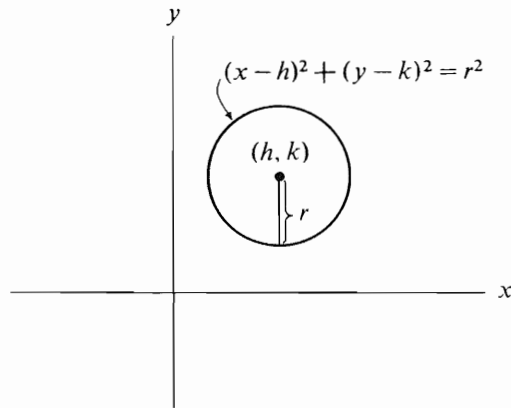
$$Ax^2 + Ay^2 + Dx + Ey + F = 0$$

(since $A = C$ and $B = 0$ in the general quadratic equation). As shown in Technical Note III, any equation that represents a circle can be written in the standard form

$$(x - h)^2 + (y - k)^2 = r^2$$

where (h, k) is the center of the circle and r is its radius, as shown in Figure 1.53.

Figure 1.53



When an attempt is made to put an equation in the standard form for a circle, degenerate or imaginary loci can be detected from the indicated value of r^2 :

If $r^2 < 0$, there is no real locus (imaginary radius)

If $r^2 = 0$, the locus is a point (zero radius)

If $r^2 > 0$, the locus is a circle

EXAMPLE

Put the following equations into the appropriate standard form for a circle, check for degenerate or imaginary loci, and graph.

- (a) $x^2 + y^2 - 4x - 6y + 19 = 0$.
- (b) $2x^2 + 2y^2 + 16x - 4y + 17 = 0$.
- (c) $x^2 + y^2 - 10x + 4y + 29 = 0$.

Standard form for a circle:

$$(x - h)^2 + (y - k)^2 = r^2$$

$$(a) \quad (x^2 - 4x + 4) + (y^2 - 6y + 9) = -19 + 4 + 9$$

$$(x - 2)^2 + (y - 3)^2 = -6$$

$r^2 < 0$, so no real locus.

$$(b) \quad 2(x^2 + 8x + 16) + 2(y^2 - 2y + 1) = -17 + 32 + 2$$

$$(x + 4)^2 + (y - 1)^2 = \frac{17}{2}$$

Circle: center $(-4, 1)$, radius $\sqrt{\frac{17}{2}}$. See Figure 1.54.

Figure 1.54

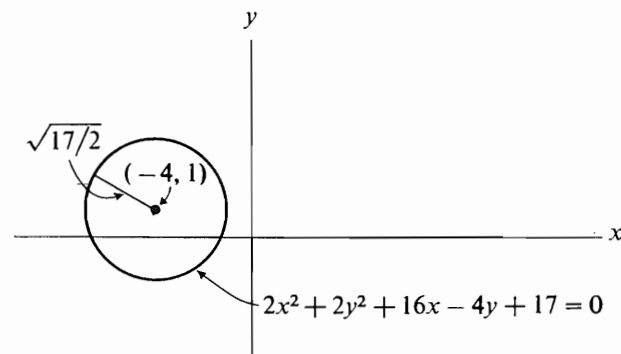
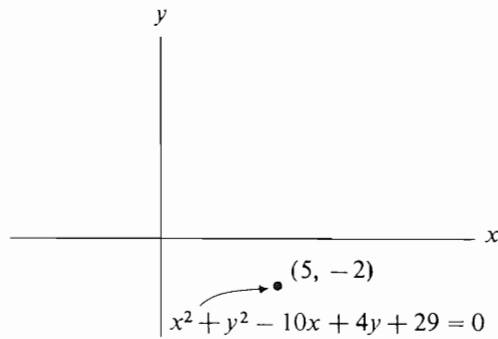


Figure 1.55



$$(c) (x^2 - 10x + 25) + (y^2 + 4y + 4) = -29 + 25 + 4$$

$$(x - 5)^2 + (y + 2)^2 = 0$$

Point: (5, -2). See Figure 1.55.

THE ELLIPSE

Geometrically, an ellipse is the locus of points in a plane the sum of whose distances from two fixed points (called *foci*) is constant. An ellipse has two (perpendicular) axes of symmetry; the longer of these is the *major axis*, the shorter is the *minor axis*; the point at which the axes intersect is the *center* of the ellipse (see Figure 1.56).

The general equation of an ellipse can be written

$$Ax^2 + Cy^2 + Dx + Ey + F = 0$$

(where $A \neq C$ and A and C are of the same sign). As shown in Technical Note III, any equation that represents an ellipse can be written in the standard form

$$\frac{(x - h)^2}{a^2} + \frac{(y - k)^2}{b^2} = 1$$

where (h, k) is the center of the ellipse and the major axis is parallel to the x -axis if $a > b$ and to the y -axis if $a < b$. (See Figure 1.57.)

The length of the axis parallel to the x -axis is $2a$ and the length of the axis parallel

Figure 1.56

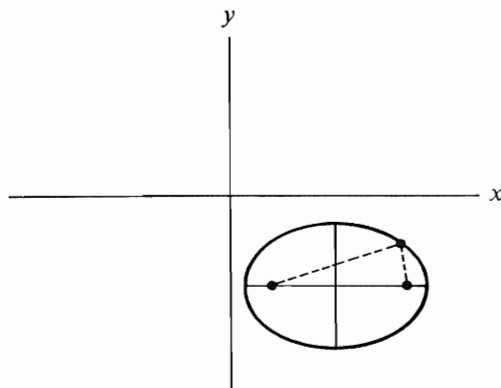
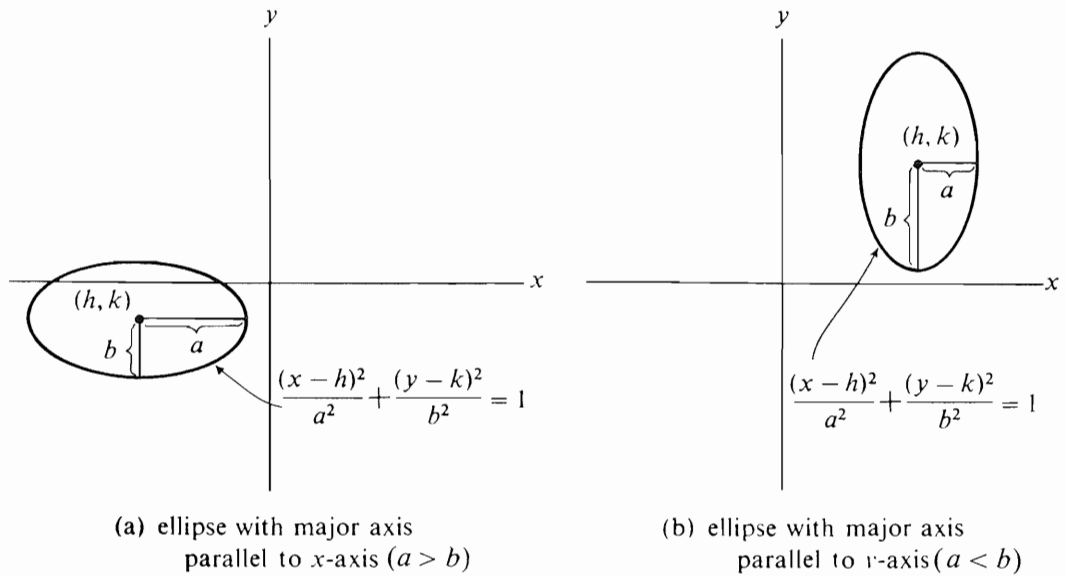


Figure 1.57



to the y-axis is $2b$. a and b are referred to as semi-axes. When $a = b$, the ellipse becomes a circle with radius $r = a = b$.

When an equation is in the standard form for an ellipse, the locus is an ellipse. However,

If $\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = C$, where $C < 0$, there is no real locus

If $\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 0$, the locus is the point (h, k)

EXAMPLE

Put the following equations into the appropriate standard form for an ellipse, check for degenerate or imaginary loci, and graph.

- (a) $6x^2 + 4y^2 - 36x + 16y + 70 = 0$.
 (b) $9x^2 + 2y^2 + 36x + 4y + 20 = 0$.
 (c) $2x^2 + y^2 - 16x - 12y + 80 = 0$.

Standard form for an ellipse:

$$\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1$$

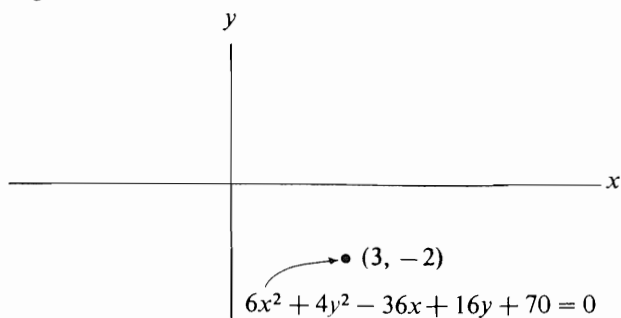
$$(a) \quad 6(x^2 - 6x + 9) + 4(y^2 + 4y + 4) = -70 + 54 + 16$$

$$6(x-3)^2 + 4(y+2)^2 = 0$$

$$\frac{(x-3)^2}{2^2} + \frac{(y+2)^2}{(\sqrt{6})^2} = 0$$

Point: $(3, -2)$. See Figure 1.58.

Figure 1.58



$$(b) \quad 9x^2 + 2y^2 + 36x + 4y + 20 = 0$$

$$9(x^2 + 4x + 4) + 2(y^2 + 2y + 1) = -20 + 36 + 2$$

$$9(x + 2)^2 + 2(y + 1)^2 = 18$$

$$\frac{(x + 2)^2}{2} + \frac{(y + 1)^2}{9} = 1$$

$$\frac{(x + 2)^2}{(\sqrt{2})^2} + \frac{(y + 1)^2}{3^2} = 1$$

Ellipse: Center $(-2, -1)$, $a = \sqrt{2}$, $b = 3$.

Major axis 6.

Minor axis $2\sqrt{2}$.

Major axis parallel to y -axis.

See Figure 1.59.

$$(c) \quad 2x^2 + y^2 - 16x - 12y + 80 = 0$$

$$2(x^2 - 8x + 16) + (y^2 - 12y + 36) = -80 + 32 + 36$$

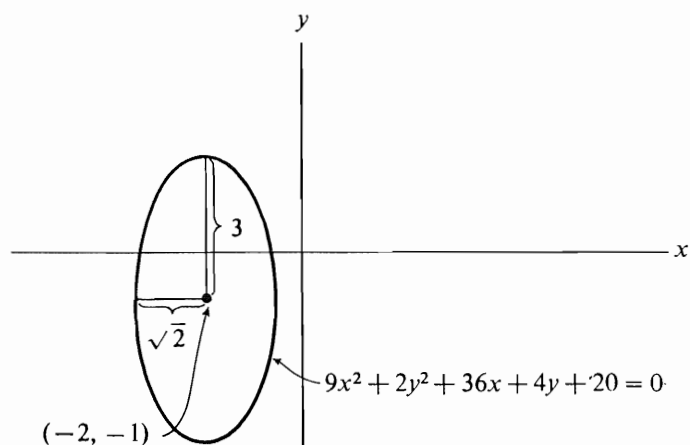
$$2(x - 4)^2 + (y - 6)^2 = -12$$

$$\frac{(x - 4)^2}{6} + \frac{(y - 6)^2}{12} = -1$$

$$\frac{(x - h)^2}{a^2} + \frac{(y - k)^2}{b^2} < 0,$$

so there is no real locus.

Figure 1.59



PROBLEMS

Determine whether each of the following equations represents a circle or an ellipse, put the equation into the appropriate standard form, check for degenerate or imaginary loci, and graph.

1. $x^2 + y^2 + 2x - 4y + 1 = 0$
2. $9x^2 + 4y^2 - 24y = 0$
3. $x^2 + 4y^2 - 6x + 16y + 45 = 0$
4. $x^2 + y^2 - 8x - 4y + 18 = 0$
5. $x^2 + y^2 - 10x + 25 = 0$
6. $x^2 + y^2 - 2x + 4y + 11 = 0$

Answers to Odd-Numbered Problems

1. circle: $(x + 1)^2 + (y - 2)^2 = 2^2$
center $(-1, 2)$
radius 2
3. ellipse: $\frac{(x - 3)^2}{4} + \frac{(y + 2)^2}{1} = -5$
so there is no real locus
5. circle: $(x - 5)^2 + (y - 0)^2 = 0$
degenerate locus: point $(5, 0)$

THE PARABOLA

Geometrically, a parabola is the locus of points in a plane which are equidistant from a fixed point (called the *focus*) and a fixed straight line (called the *directrix*). A parabola is symmetric about a line, called its *axis*; the point of intersection of the parabola and its axis is the *vertex* of the parabola (see Figure 1.60).

The general equation of a parabola can be written

$$Ax^2 + Dx + Ey + F = 0$$

if the axis is parallel to the y -axis and

$$Cy^2 + Dx + Ey + F = 0$$

if the axis is parallel to the x -axis. As shown in Technical Note III, any equation that represents a parabola can be written in the standard form

$$(x - h)^2 = 4p(y - k)$$

Figure 1.60

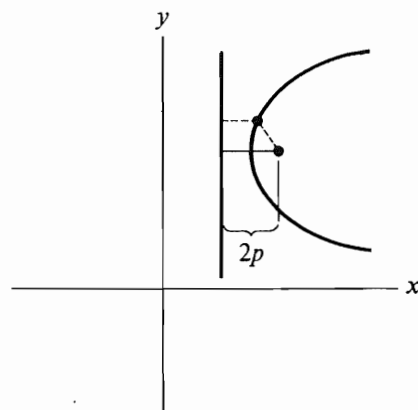
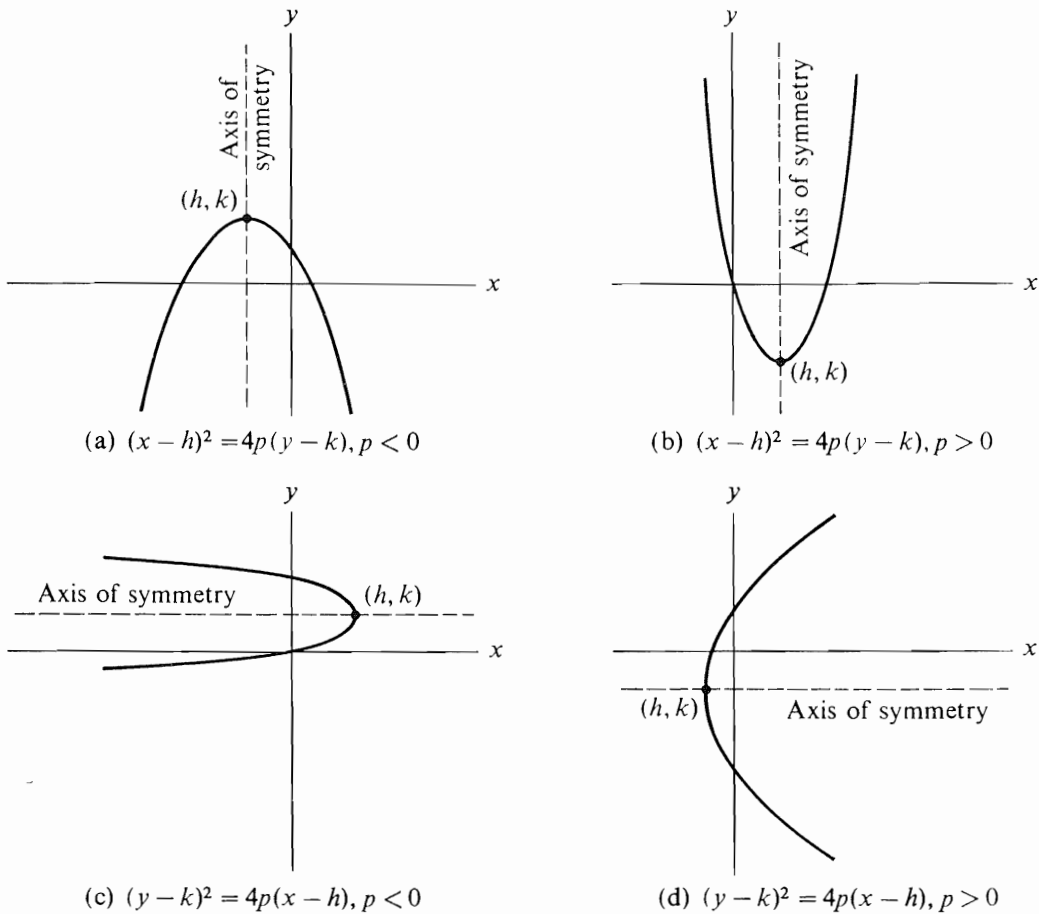


Figure 1.61



where (h, k) is the vertex of the parabola and the axis is parallel to the y -axis, or in the standard form

$$(y - k)^2 = 4p(x - h)$$

where (h, k) is the vertex of the parabola and the axis is parallel to the x -axis. The orientation and curvature of a parabola are determined by the sign and magnitude of p as follows:

For a parabola with axis parallel to the y -axis:

If $p < 0$, the parabola opens downward

If $p > 0$, the parabola opens upward

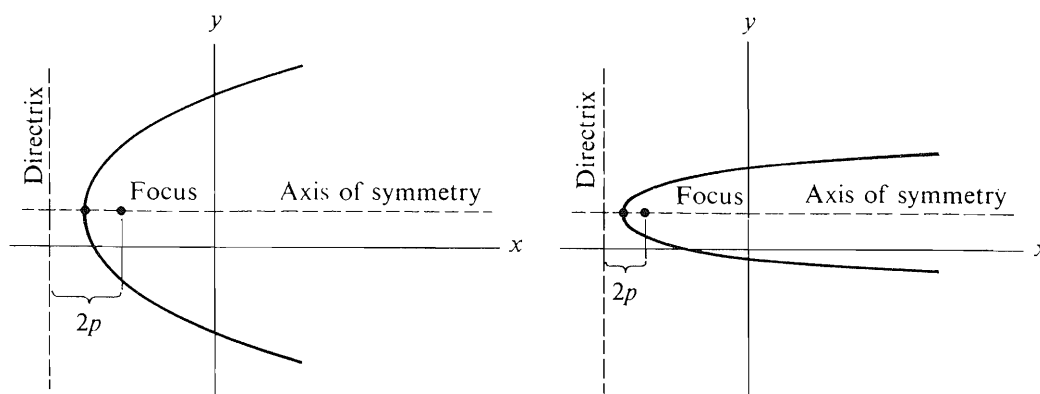
For a parabola with axis parallel to the x -axis:

If $p < 0$, the parabola opens to the left

If $p > 0$, the parabola opens to the right

(see Figure 1.61). The distance between the point (focus) and the line (directrix) defining a parabola is $2p$ (distance from a point to a line is measured along the perpendicular); as shown in Figure 1.62, the larger the value of p , the faster the parabola opens.

Figure 1.62



When an equation is in either standard form for a parabola, the locus is a parabola. However,

If $(x - h)^2 = C$, where $C < 0$, there is no real locus

If $(x - h)^2 = 0$, the locus is two coincident straight lines
(that is, geometrically, one straight line)

If $(x - h)(x - k) = 0$, the locus is two parallel lines

(and similarly for a parabola with axis parallel to the x -axis).

EXAMPLE

Put the following equations into the appropriate standard form for a parabola, check for degenerate or imaginary loci, and graph.

- (a) $y^2 - 8y + 17 = 0$.
- (b) $x^2 - 4x + y + 14 = 0$.
- (c) $x^2 - 7x + 12 = 0$.
- (d) $y^2 - 10y + 25 = 0$.
- (e) $y^2 - 2y - 6x + 19 = 0$.

Standard form for a parabola:

$$(y - k)^2 = 4p(x - h) \quad \text{or} \quad (x - h)^2 = 4p(y - k)$$

$$(a) \quad y^2 - 8y + 17 = 0$$

$$(y^2 - 8y + 16) = -17 + 16$$

$$(y - 4)^2 = -1$$

$(y - k)^2 < 0$, so there is no real locus. Note that the x -term is missing in this example.

$$(b) \quad x^2 - 4x + y + 14 = 0$$

$$x^2 - 4x + 4 = -y - 14 + 4$$

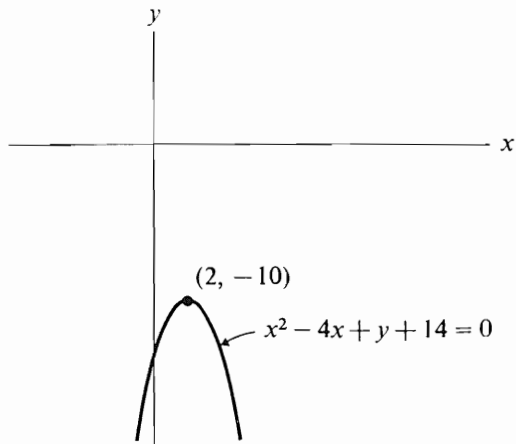
$$(x - 2)^2 = -(y + 10)$$

Parabola: Vertex $(2, -10)$, axis parallel to the y -axis, parabola opens downward.

See Figure 1.63. (Note that $4p = -1$.)

Note the distinction between example (a), where the locus is imaginary, and example (b), where the locus is real. In (a), the square of a number is said to be a negative constant, which is

Figure 1.63



not possible for a real number. In (b), the square of a number is said to be the negative of a quantity involving a variable, which implies that the quantity $(y + 10)$ must be negative or zero; $y + 10$ is negative or zero if y is limited in extent to the interval $y \leq -10$, as shown in Figure 1.63.

$$(c) \quad x^2 - 7x + 12 = 0$$

$$x^2 - 7x + \frac{49}{4} = -12 + \frac{49}{4}$$

$$(x - \frac{7}{2})^2 = \frac{1}{4}$$

$$x - \frac{7}{2} = \pm \sqrt{\frac{1}{4}}$$

$$x - \frac{7}{2} = \pm \frac{1}{2}$$

$$x = \frac{8}{2} \quad \text{or} \quad x - 4 = 0$$

and

$$x = \frac{6}{2} \quad \text{or} \quad x - 3 = 0$$

The locus is two parallel straight lines, parallel to the y -axis (see Figure 1.64). Note that the y -term is missing in this example and the equation could have been factored $(x - 4) \times (x - 3) = 0$ without completing the square.

$$(d) \quad y^2 - 10y + 25 = 0$$

$$(y - 5)^2 = 0$$

Figure 1.64

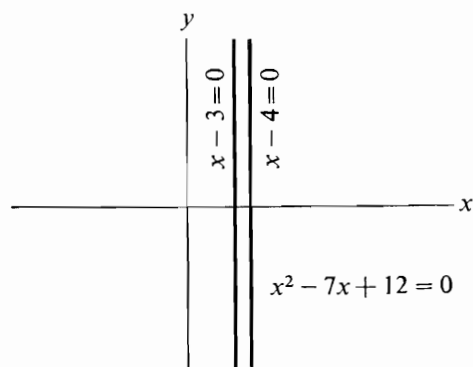
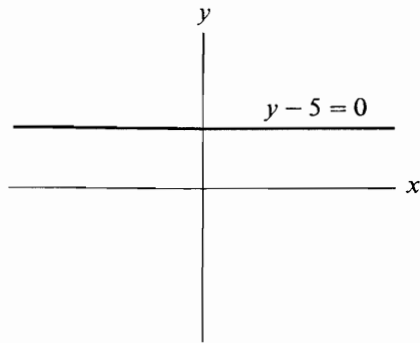


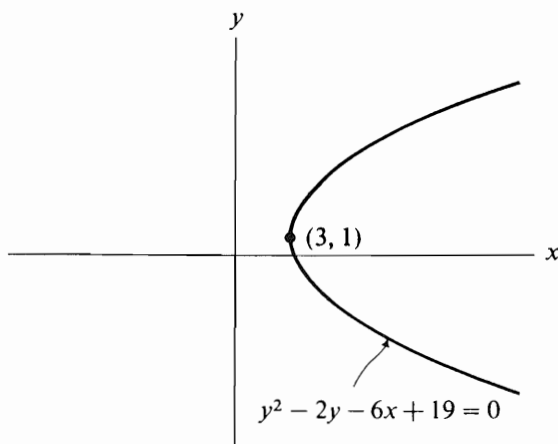
Figure 1.65



The locus is two coincident straight lines (see Figure 1.65). Note again that the x -term is missing.

$$\begin{aligned}
 \text{(e) } y^2 - 2y - 6x + 19 &= 0 \\
 (y^2 - 2y + 1) &= 6x - 19 + 1 \\
 (y - 1)^2 &= 6x - 18 \\
 (y - 1)^2 &= 6(x - 3)
 \end{aligned}$$

Figure 1.66



Parabola: Vertex $(3, 1)$, axis parallel to the x -axis, parabola opens to the right (see Figure 1.66).

THE HYPERBOLA

Geometrically, a hyperbola is the locus of points in a plane the difference of whose distances from two fixed points (called *foci*) is a constant. A hyperbola has two (perpendicular) axes of symmetry; the axis that intersects the hyperbola is its *transverse axis*; the point at which the axes intersect is the *center* of the hyperbola. A hyperbola always has two branches.

The general equation of a hyperbola can be written

$$Ax^2 + Cy^2 + Dx + Ey + F = 0$$

(where A and C are opposite in sign.) As shown in Technical Note III, any equation that represents a hyperbola can be written in the standard form

$$\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = 1$$

where (h, k) is the center of the hyperbola and the transverse axis is parallel to the x -axis, or in the standard form

$$\frac{(y-k)^2}{b^2} - \frac{(x-h)^2}{a^2} = 1$$

where (h, k) is the center of the hyperbola and the transverse axis is parallel to the y -axis. (See Figure 1.67.)

Every hyperbola has a pair of intersecting straight lines as asymptotes; these asymptotes are given by the equations

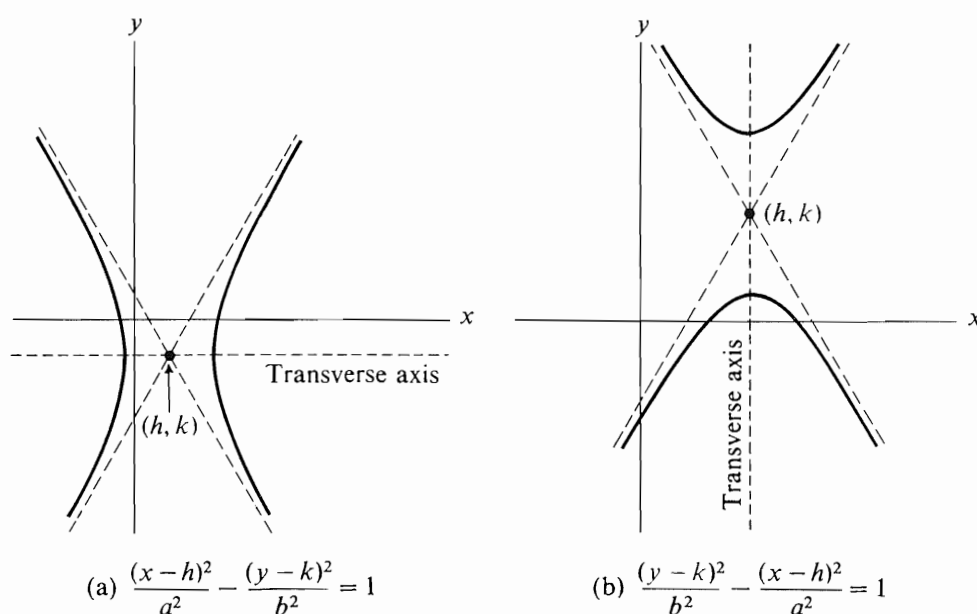
$$\frac{x-h}{a} = \pm \frac{y-k}{b}$$

Note that the equations for these asymptotes can be obtained from the standard form by replacing 1 by 0, transposing the negative term, and taking square roots.

If $a = b$, the asymptotes of the hyperbola are perpendicular and the hyperbola is a *rectangular* or *equilateral hyperbola*. Equilateral hyperbolas are discussed in more detail below.

When an equation is in the standard form for a hyperbola, the locus is a hyperbola. However,

Figure 1.67



If $\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = 0$, the locus is two intersecting straight lines,

$$\frac{x-h}{a} = \pm \frac{y-k}{b}$$

Note that these are also the equations for the asymptotes of the hyperbola.

There is no possibility of an imaginary locus in the case of a hyperbola. If, for example,

$$\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = C \quad \text{where } C < 0$$

then

$$\frac{(y-k)^2}{b^2} - \frac{(x-h)^2}{a^2} = C \quad \text{where } C > 0$$

EXAMPLE

Put the following equations into the appropriate standard form for a hyperbola, check for degenerate loci, and graph.

(a) $6x^2 - 12x - 4y^2 - 16y - 34 = 0$.

(b) $2y^2 - 12y - x^2 + 6x + 7 = 0$.

(c) $5x^2 + 20x - 3y^2 - 24y - 28 = 0$.

Standard form for a hyperbola:

$$\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = 1 \quad \text{or} \quad \frac{(y-k)^2}{b^2} - \frac{(x-h)^2}{a^2} = 1$$

(a) $6x^2 - 12x - 4y^2 - 16y - 34 = 0$

$$(6x^2 - 12x) - (4y^2 + 16y) = 34$$

$$6(x^2 - 2x + 1) - 4(y^2 + 4y + 4) = 34 + 6 - 16$$

$$6(x-1)^2 - 4(y+2)^2 = 24$$

$$\frac{(x-1)^2}{4} - \frac{(y+2)^2}{6} = 1$$

$$\frac{(x-1)^2}{2^2} - \frac{(y+2)^2}{(\sqrt{6})^2} = 1$$

Hyperbola: Center $(1, -2)$, $a = 2$, $b = \sqrt{6}$, transverse axis parallel to the x -axis.

Asymptotes: $\frac{x-1}{2} = \pm \frac{y+2}{\sqrt{6}}$

$$y+2 = \pm \frac{\sqrt{6}}{2}(x-1)$$

Plotting the asymptotes:

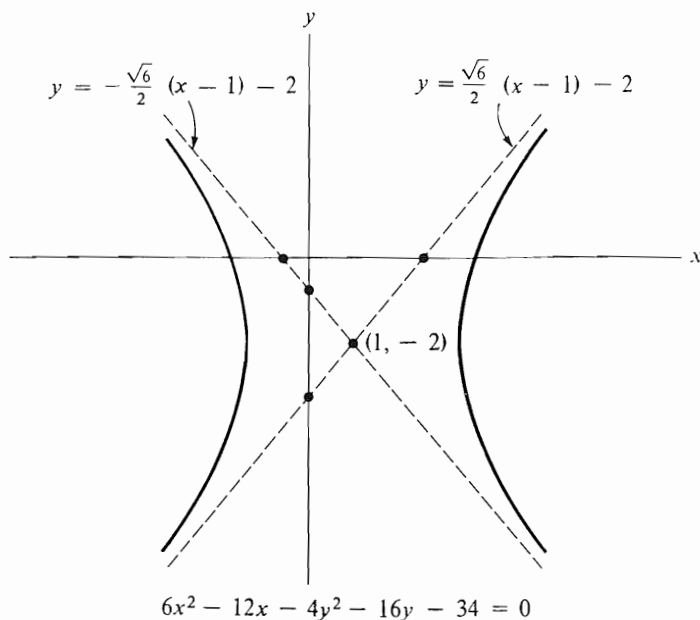
$$y \approx 1.23(x-1) - 2 \quad y \approx -1.23(x-1) - 2$$

$$\text{If } x = 0, y \approx -3.23 \quad \text{If } x = 0, y \approx -0.77$$

$$\text{If } y = 0, x \approx 2.63 \quad \text{If } y = 0, x \approx -0.63$$

(see Figure 1.68).

Figure 1.68



NOTE: The symbol \approx means "approximately equal to" and is used because of the approximation of $\sqrt{6}$.

$$\begin{aligned}
 \text{(b)} \quad & 2y^2 - 12y - x^2 + 6x + 7 = 0 \\
 & 2(y^2 - 6y + 9) - (x^2 - 6x + 9) = -7 + 18 - 9 \\
 & 2(y-3)^2 - (x-3)^2 = 2 \\
 & \frac{(y-3)^2}{1} - \frac{(x-3)^2}{2} = 1 \\
 & \frac{(y-3)^2}{1^2} - \frac{(x-3)^2}{(\sqrt{2})^2} = 1
 \end{aligned}$$

Hyperbola: Center $(3, 3)$; $a = \sqrt{2}$, $b = 1$, transverse axis parallel to the y -axis.

Asymptotes: $y - 3 = \pm \frac{1}{\sqrt{2}}(x - 3)$

Plotting the asymptotes:

$$y \approx 0.71(x - 3) + 3 \quad y \approx -0.71(x - 3) + 3$$

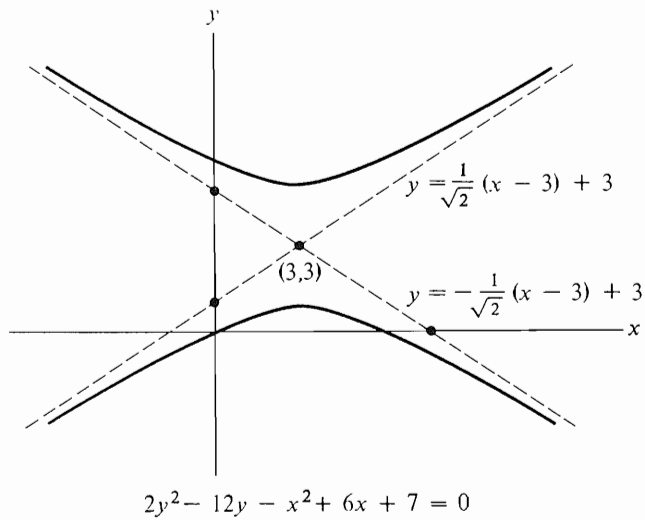
$$\text{If } x = 0, y \approx 0.87 \quad \text{If } x = 0, y \approx 5.13$$

$$\text{If } y = 0, x \approx -1.23 \quad \text{If } y = 0, x \approx 7.23$$

See Figure 1.69.

$$\begin{aligned}
 \text{(c)} \quad & 5x^2 + 20x - 3y^2 - 24y - 28 = 0 \\
 & 5(x^2 + 4x + 4) - 3(y^2 + 8y + 16) = 28 + 20 - 48 \\
 & 5(x+2)^2 - 3(y+4)^2 = 0 \\
 & \frac{(x+2)^2}{3} - \frac{(y+4)^2}{5} = 0
 \end{aligned}$$

Figure 1.69



The locus is two straight lines,

$$y = \pm \sqrt{\frac{5}{3}}(x + 2) - 4$$

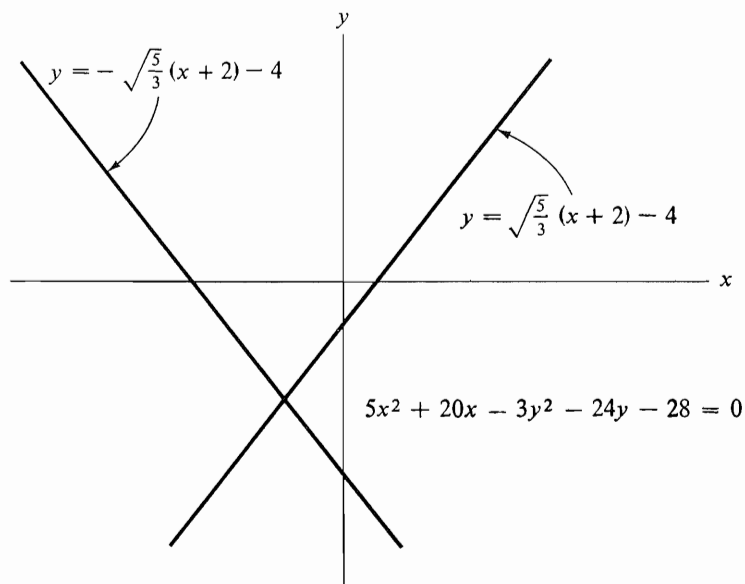
$$y \approx 1.29(x + 2) - 4 \quad y \approx -1.29(x + 2) - 4$$

$$\text{If } x = 0, y \approx -1.42 \quad \text{If } x = 0, y \approx -6.58$$

$$\text{If } y = 0, x \approx 1.10 \quad \text{If } y = 0, x \approx -5.10$$

See Figure 1.70.

Figure 1.70



Special Cases of the Equilateral Hyperbola

In the preceding discussion of the hyperbola, an equilateral hyperbola was defined as a hyperbola whose asymptotes are perpendicular to each other. A hyperbola whose equation can be written in the standard form

$$\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = 1 \quad \text{or} \quad \frac{(y-k)^2}{b^2} - \frac{(x-h)^2}{a^2} = 1$$

is an equilateral hyperbola if and only if $a = b$.

The equation of an equilateral hyperbola whose asymptotes are parallel, respectively, to the coordinate axes can be written in the standard form

$$(x-h)(y-k) = c$$

where (h, k) is the center of the hyperbola, about which it is symmetric, and $x = h$ and $y = k$ are the asymptotes. (Note that this is the special case $A = C = 0, B \neq 0$, of the general second-degree equation $Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$.) If the asymptotes are considered as coordinate axes, the hyperbola has its two branches in the first and third quadrants if $c > 0$ and in the second and fourth quadrants if $c < 0$, as shown in Figure 1.71. The intercepts of the hyperbola are the points $(h - (c/k), 0)$ and $(0, k - (c/h))$, where $h - (c/k)$ and $k - (c/h)$ may be positive or negative according to the values of h, k , and c . Note that this type of hyperbola has no more than one x -intercept and one y -intercept; for a hyperbola to have two x -intercepts, its equation must include an x^2 -term, and for a hyperbola to have two y -intercepts, its equation must include a y^2 -term. Similarly, for a hyperbola to have one y -intercept, its equation must include an x -term and for a hyperbola to have one x -intercept, its equation must include a y -term.

The equation of an equilateral hyperbola whose asymptotes are the coordinate axes can be written in the standard form

$$xy = c$$

[Note that this is the special case $h = k = 0$ of the equation $(x-h)(y-k) = c$.]

Figure 1.71

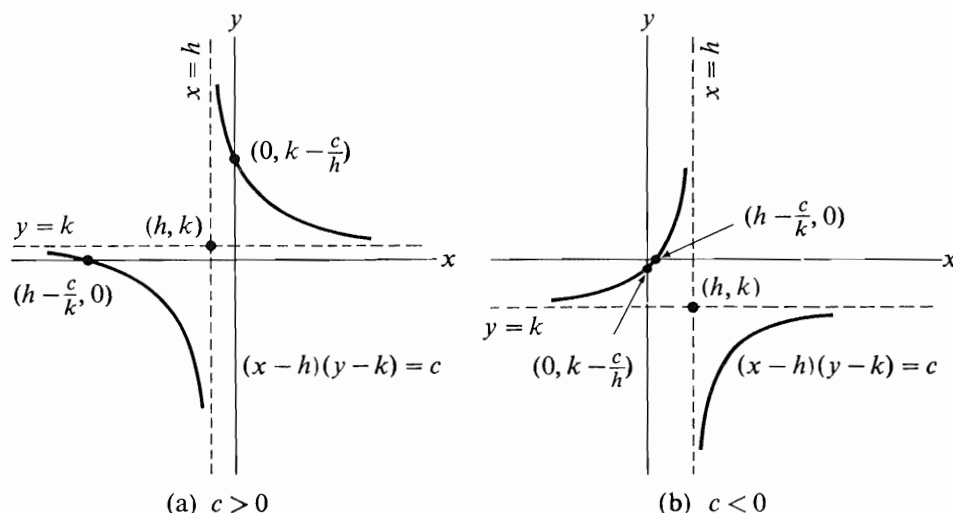
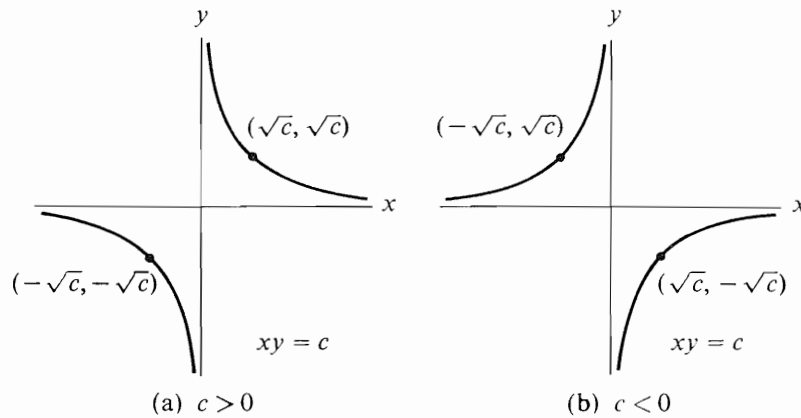


Figure 1.72



This type of hyperbola has its center at the origin, about which it is symmetric; its branches occur in the first and third quadrants if $c > 0$ and in the second and fourth quadrants if $c < 0$. Since it is asymptotic to both coordinate axes, the hyperbola has no intercepts (see Figure 1.72).

The equation $xy = c$ expresses an inversely proportional relationship between x and y —that is, as one variable increases, the other variable decreases proportionally. The variable y is said to be *inversely proportional* to the variable x , if there is a constant c such that

$$y = \frac{c}{x} \quad \text{or} \quad xy = c$$

By a generalization of this definition, the variable y is inversely proportional to a positive power n of a variable x , if there is a constant c , such that

$$y = \frac{c}{x^n} \quad \text{or} \quad x^n y = c$$

The curve corresponding to such an equation or to the analogous equation $xy^m = c$ is a *generalized equilateral hyperbola*, or *hyperbola of Fermat*. Its center is at the origin; its asymptotes are the coordinate axes.

If n or m is odd, the branches of the equilateral hyperbola are symmetric to the origin and are in the first and third quadrants if $c > 0$; the branches are in the second and fourth quadrants if $c < 0$. If $x^n y = c$ and n is even, the branches of the equilateral hyperbola are symmetric to the y -axis and are in the first and second quadrants if $c > 0$ and are in the third and fourth quadrants if $c < 0$. If $xy^m = c$ and m is even, the branches of the equilateral hyperbola are symmetric to the x -axis and are in the first and fourth quadrants if $c > 0$ and are in the second and third quadrants if $c < 0$.

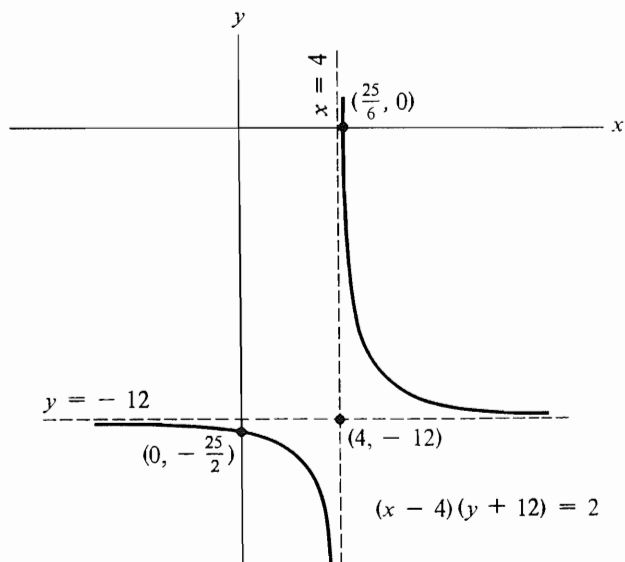
EXAMPLE

Graph the equation

$$(x - 4)(y + 12) = 2$$

(see Figure 1.73).

Figure 1.73



Equilateral hyperbola: Center: $(4, -12)$.

Asymptotes: $x = 4$, $y = -12$.

Lies in the first and third quadrants of its asymptotes.

EXAMPLE

Graph the equation

$$(x - 2)y = -4$$

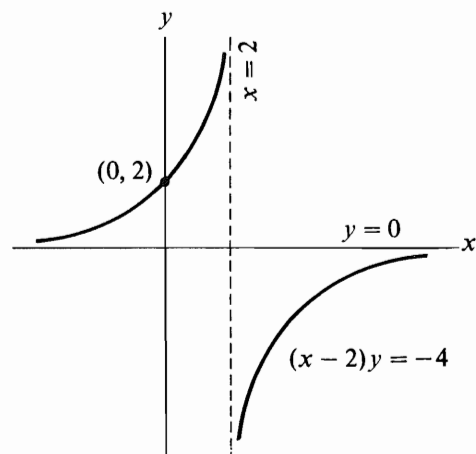
(see Figure 1.74).

Equilateral hyperbola: Center: $(2, 0)$.

Asymptotes: $x = 2$, $y = 0$.

Lies in the second and fourth quadrants of its asymptotes.

Figure 1.74



EXAMPLE

Graph the equation

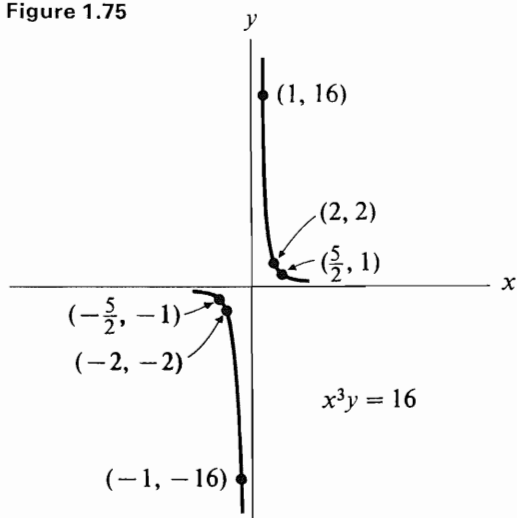
$$x^3y = 16$$

(see Figure 1.75).

Generalized equilateral hyperbola, asymptotic to coordinate axes, symmetric to origin.

Points to aid graphing: $(2, 2)$, $(1, 16)$, $(\sqrt[3]{16}, 1)$, $(-2, -2)$, $(-1, -16)$, $(-\sqrt[3]{16}, -1)$ (note that $\sqrt[3]{16} \approx \frac{5}{2}$).

Figure 1.75



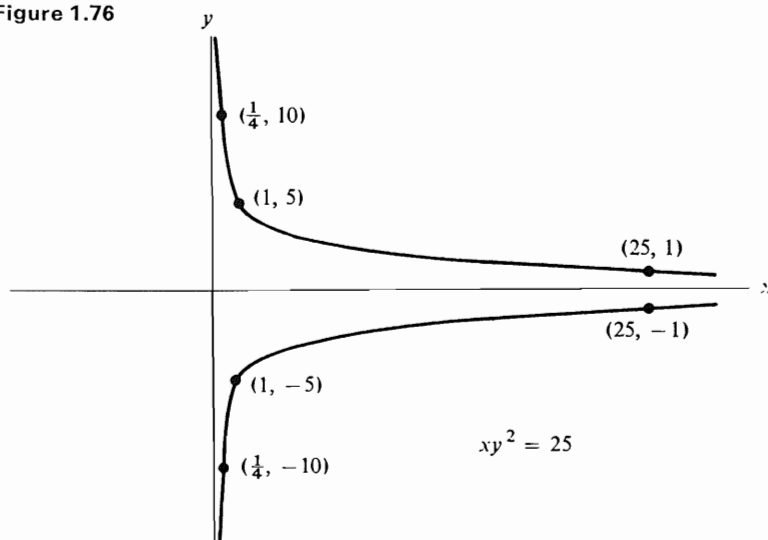
EXAMPLE

Graph the equation

$$xy^2 = 25$$

(See Figure 1.76.)

Figure 1.76



Generalized equilateral hyperbola, asymptotic to coordinate axes, symmetric to x -axis.

Points to aid graphing: $(1, 5)$, $(25, 1)$, $(\frac{1}{4}, 10)$, $(1, -5)$, $(25, -1)$, $(\frac{1}{4}, -10)$.

PROBLEMS

Determine whether each of the following equations represents a parabola or a hyperbola, put the equation into the appropriate standard form, check for degenerate or imaginary loci, and graph.

- | | |
|-------------------------------------|---------------------------|
| 1. $y^2 - 2y - 2x + 9 = 0$ | 4. $y^2 - 8y + 24 = 0$ |
| 2. $x^2 - 3y^2 - 4x + 12y - 11 = 0$ | 5. $xy - 4x - 5y + 5 = 0$ |
| 3. $3x^2 - 2y^2 - 6x - 4y + 1 = 0$ | 6. $xy + 5x - y - 5 = 0$ |

Answers to Odd-Numbered Problems

- | | |
|--|--|
| 1. parabola: $(y - 1)^2 = 2(x - 4)$
vertex $(4, 1)$
axis parallel to x -axis, opens to right | 5. equilateral hyperbola: $(x - 5)(y - 4) = 15$
center $(5, 4)$
asymptotes $x = 5$, $y = 4$
first and third axes quadrants |
| 3. degenerate hyperbola: two intersecting
lines $3x - 2y - 4 = 0$ and $3x + 2y - 2 = 0$ | |

FAMILIES OF QUADRATIC CURVES

The general equation of the second degree

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$$

has five arbitrary constants, since if any of the constants A , B , C , D , E , and F is not zero, it may be used as a divisor to reduce the equation to a form involving only five arbitrary constants; after the values of five constants are set arbitrarily, the value of the other constant must be set so that the sum of the terms is zero. These five arbitrary constants in the equation of a second-degree curve permit five geometric conditions to be imposed on the curve.

As discussed above, various conditions on the constants A , B , and C determine the general form of the second-degree curve; complete specification of the constants determines a particular curve. Families of various types of curves are represented by an equation in which one (or more) of the constants is specified and the others are unspecified (that is, remain parameters). For example, a family of circles may have the same center (that is, be concentric) but have different radii, or a family of parabolas may have the same vertex but have different degrees of curvature. Such families of curves will be discussed in more detail in later sections.

PROBLEMS

For each of the following equations, identify the curve represented, put the equation into the appropriate standard form and identify the parameters and properties thus obtained, and sketch the curve.

- | | |
|----------------------------------|----------------------------------|
| 1. $x^2 + y^2 - 6x - 2y - 6 = 0$ | 4. $y^2 - 10y = 0$ |
| 2. $y^2 - 6y + 9 = 0$ | 5. $xy - 4y = -4$ |
| 3. $3x^2 + 3y^2 - 6x + 4y = 1$ | 6. $x^2 - y^2 + 4x - 2y + 1 = 0$ |

7. $2x^2 + y^2 = 50$
8. $x^2 + y^2 - 4x - 2y + 5 = 0$
9. $4x^2 + 9y^2 - 16x - 18y + 133 = 0$
10. $xy + 3y = x + 6$
11. $3x^2 - y^2 - 12x - 6y = 0$
12. $x^2 - y^2 - 16 = 0$
13. $y = 3 + 2x - x^2$
14. $9x^2 + 25y^2 + 18x + 150y + 9 = 0$
15. $x^2 + 9y^2 - 8x + 7 = 0$
16. $16x^2 + y^2 - 32x - 6y + 25 = 0$
17. $y^2 - 3x^2 = 27$
18. $2x = 5y - y^2$
19. $5x^2 + 4y = 12$
20. $xy + 15y + 3x = 15$
21. $y^2 - 2y - 8x + 25 = 0$
22. $x^2 + y^2 - 4x - 2y + 6 = 0$
23. $y^2 - 4x^2 - 4y + 4 = 0$
24. $y^2 - 12y + 46 = 0$
25. $3y^2 + 2x = 0$
26. $xy - 6x + 2 = 0$

Answers to Odd-Numbered Problems

1. circle: $(x - 3)^2 + (y - 1)^2 = 4^2$
center (3, 1)
radius 4
3. circle: $(x - 1)^2 + (y + \frac{2}{3})^2 = (\frac{4}{3})^2$
center $(1, -\frac{2}{3})$
radius $\frac{4}{3}$
5. equilateral hyperbola: $y(x - 4) = -4$
center (4, 0)
asymptotes $x = 4, y = 0$
second and fourth axes quadrants
7. ellipse: $\frac{(x - 0)^2}{5^2} + \frac{(y - 0)^2}{(5\sqrt{2})^2} = 1$
center (0, 0)
semiaxes $5\sqrt{2}, 5$
major axis parallel to y -axis
9. ellipse with no real locus: $\frac{(x - 2)^2}{3^2} + \frac{(y - 1)^2}{2^2} = -3$
11. hyperbola: $\frac{(x - 2)^2}{1^2} - \frac{(y + 3)^2}{(\sqrt{3})^2} = 1$
center (2, -3)
asymptotes $y = \sqrt{3}x - 2\sqrt{3} - 3$ and $y = -\sqrt{3}x + 2\sqrt{3} - 3$
transverse axis parallel to x -axis
13. parabola: $(x - 1)^2 = -(y - 4)$
vertex (1, 4)
axis parallel to y -axis, opens downward
15. ellipse: $\frac{(x - 4)^2}{3^2} + \frac{(y - 0)^2}{1^2} = 1$
center (4, 0)
semiaxes 3, 1
major axis parallel to x -axis
17. hyperbola: $\frac{(y - 0)^2}{(3\sqrt{3})^2} - \frac{(x - 0)^2}{3^2} = 1$
center (0, 0)
asymptotes $y = \sqrt{3}x$ and $y = -\sqrt{3}x$
transverse axis parallel to y -axis

19. parabola: $(x - 0)^2 = -\frac{4}{3}(y - 3)$
 vertex (0, 3)
 axis parallel to y -axis, opens downward
21. parabola: $(y - 1)^2 = 8(x - 3)$
 vertex (3, 1)
 axis parallel to x -axis, opens to right
23. degenerate hyperbola: two intersecting lines: $y = 2x + 2$ and
 $y = -2x + 2$
25. parabola: $(y - 0)^2 = -\frac{2}{3}(x - 0)$
 vertex (0, 0)
 axis parallel to x -axis, opens to left

1.7 Applications of Quadratic Curves in Business and Economics

In this section several applications of quadratic curves in business and economics are discussed. These applications include quadratic demand and supply curves and the associated market equilibria, product transformation curves, and Pareto's law of the distribution of incomes.

DEMAND AND SUPPLY CURVES

The first quadrant parts of various types of parabolas frequently are appropriate for representing demand and supply functions, as illustrated in Figures 1.77 and 1.78. The first quadrant part of an equilateral hyperbola is frequently used to represent a demand function, as illustrated in Figure 1.79.

Note that each curve in Figures 1.77, 1.78, and 1.79 is only one of a family of curves appropriate for representing the functions discussed. For example, the vertex of the parabola in Figure 1.77(a) can lie anywhere in the second quadrant or on the

Figure 1.77 Examples of parabolic demand functions.

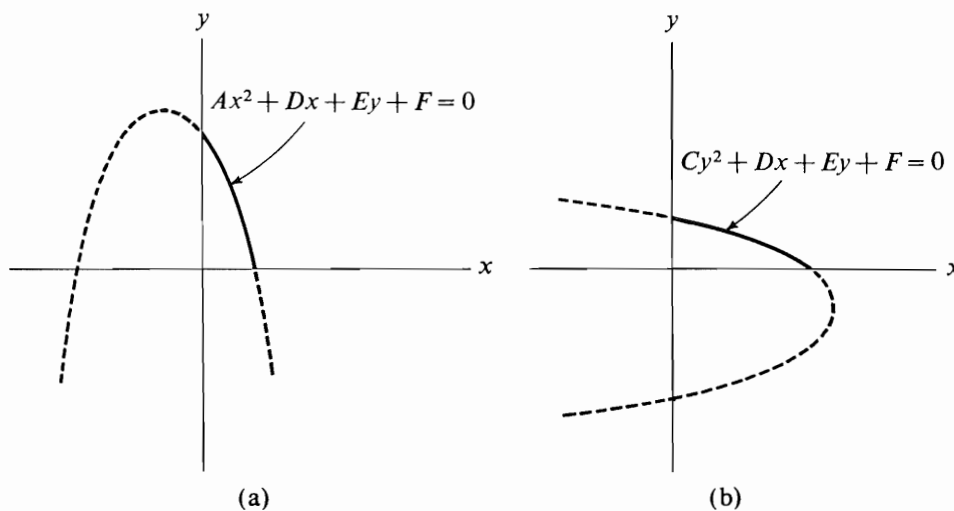
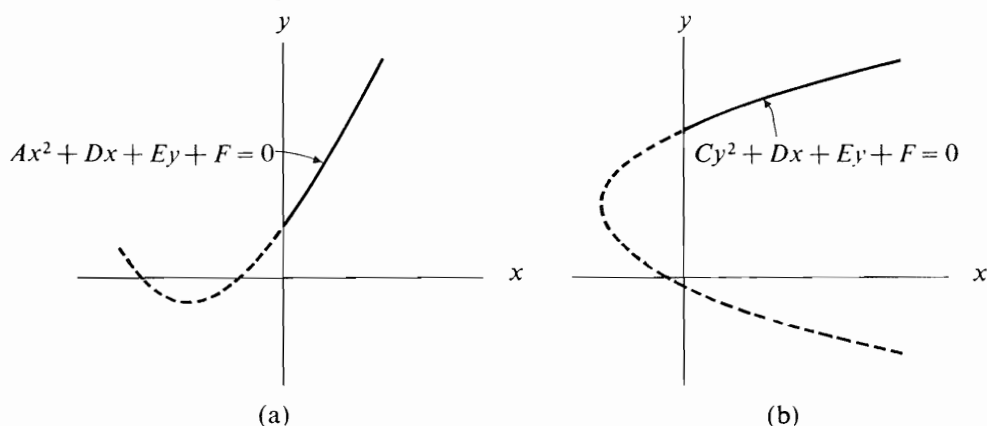
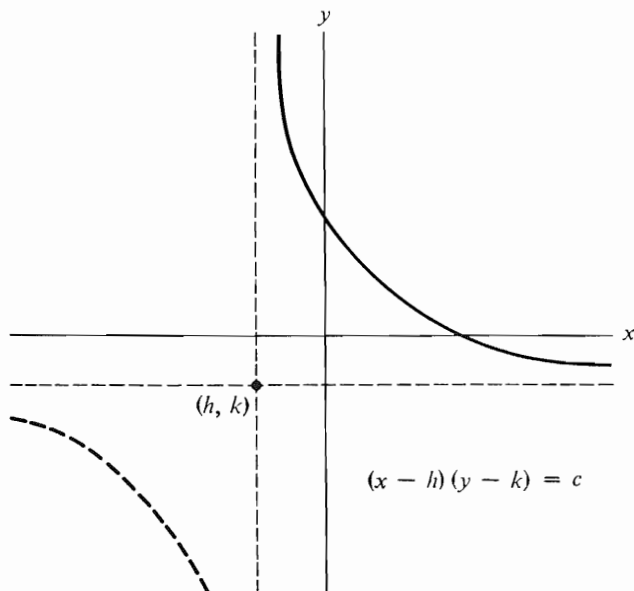


Figure 1.78 Examples of parabolic supply functions.



positive y -axis, provided the parabola has positive x - and y -intercepts. The conditions on the coefficients of the equation which specify the family of curves appropriate for representing a particular type of economic function can be determined using the corresponding standard form of the equation (see the problems).

Figure 1.79 Example of a hyperbolic demand function.



Market Equilibrium

Market equilibrium price and quantity are the price and quantity represented by the coordinates of the point of intersection of the demand and supply curves. An approximate solution for these coordinates can be obtained geometrically for any supply and demand curves. However, algebraic simultaneous solution, even for second-degree demand and supply functions, may involve solution of a third- or fourth-degree equation.

Since general methods for solution of equations of degree greater than two are not discussed here, the examples and problems are limited to cases involving solution of

a quadratic equation. For example, if one equation is linear and the other equation is parabolic or hyperbolic, their simultaneous solution involves solution of a quadratic equation; similarly, simultaneous solution of two parabolic equations which are quadratic with respect to the same variable involves solution of a quadratic equation.

Unless a quadratic equation can be factored, its solution is most readily obtained by using the quadratic formula: The root(s) of a quadratic equation $ax^2 + bx + c = 0$ are given by

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

This formula is obtained by the process of completing the square, as shown in Technical Note IV. Note that the number of real roots an equation has is determined by the value of $b^2 - 4ac$:

If $b^2 - 4ac < 0$, no real roots

If $b^2 - 4ac = 0$, one real root

If $b^2 - 4ac > 0$, two real roots

EXAMPLE

Find the equilibrium price and quantity for the following demand and supply equations (where x represents quantity and y represents price):

$$2x + y - 10 = 0$$

$$y^2 - 8x - 4 = 0$$

Solving the equations simultaneously,

$$2x = -y + 10 \quad \text{from the first equation}$$

$$8x = y^2 - 4 \quad \text{from the second equation}$$

so

$$y^2 - 4 = -4y + 40$$

$$y^2 + 4y - 44 = 0$$

and, using the quadratic formula,

$$y = \frac{-4 \pm \sqrt{16 - (4)(-44)}}{2}$$

$$= \frac{-4 \pm 8\sqrt{3}}{2}$$

$$= -2 \pm 4\sqrt{3}$$

$$x = \frac{1}{2}(-y + 10)$$

$$= 6 \pm 2\sqrt{3}$$

$\sqrt{3} \approx 1.732$, so the approximate solutions are (2.5, 4.9) and (9.5, -8.9) and the approximate equilibrium point is (2.5, 4.9).

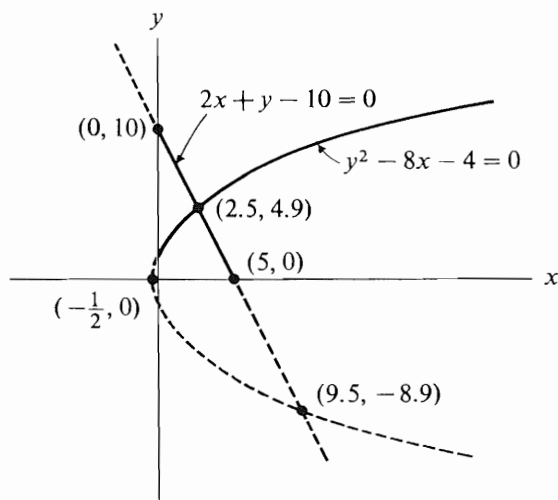
Plotting the curves to verify the result geometrically (see Figure 1.80),

$$2x + y - 10 = 0 \quad y^2 - 8x - 4 = 0$$

$$\text{If } x = 0, y = 10 \quad y^2 = 8x + 4$$

$$\text{If } y = 0, x = 5 \quad y^2 = 8(x + \frac{1}{2})$$

Figure 1.80



EXAMPLE

Find the equilibrium price and quantity for the following demand and supply equations (where x represents quantity and y represents price):

$$x^2 + 5x - y + 1 = 0$$

$$2x^2 + y - 9 = 0$$

Solving the equations simultaneously,

$$y = x^2 + 5x + 1$$

$$y = -2x^2 + 9$$

so

$$x^2 + 5x + 1 = -2x^2 + 9$$

$$3x^2 + 5x - 8 = 0$$

$$(3x + 8)(x - 1) = 0$$

$$x = -\frac{8}{3}, 1$$

$$y = -2x^2 + 9$$

$$y = -\frac{47}{9}, 7$$

The solutions are $(-\frac{8}{3}, -\frac{47}{9})$ and $(1, 7)$ and the equilibrium point is $(1, 7)$.

Plotting the curves to verify the result geometrically (see Figure 1.81),

$$x^2 + 5x - y + 1 = 0$$

$$2x^2 + y - 9 = 0$$

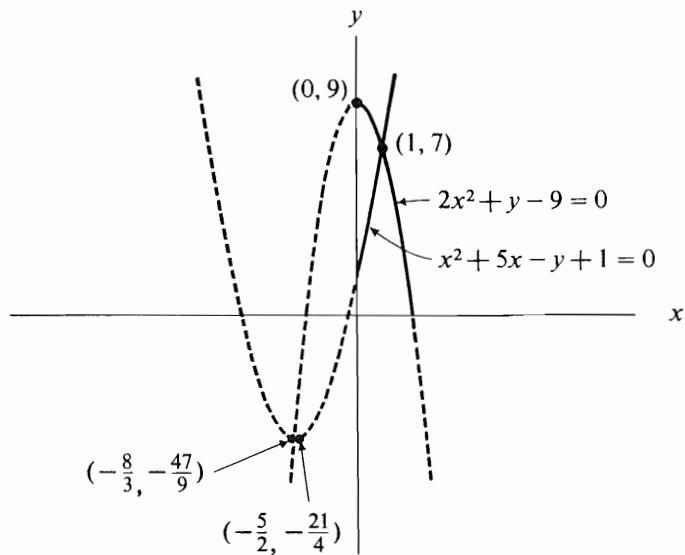
$$(x^2 + 5x + \frac{25}{4}) = y - 1 + \frac{25}{4}$$

$$2x^2 = -(y - 9)$$

$$(x + \frac{5}{2})^2 = y + \frac{21}{4}$$

$$x^2 = -\frac{1}{2}(y - 9)$$

Figure 1.81

**EXAMPLE**

Find the equilibrium price and quantity for the following demand and supply equations (where x represents quantity and y represents price):

$$y^2 + y + x - 20 = 0$$

$$2y^2 - x - 3y - 4 = 0$$

Solving the equations simultaneously,

$$x = -y^2 - y + 20$$

$$x = 2y^2 - 3y - 4$$

so

$$-y^2 - y + 20 = 2y^2 - 3y - 4$$

$$3y^2 - 2y - 24 = 0$$

and, using the quadratic formula,

$$y = \frac{2 \pm \sqrt{4 - 4(3)(-24)}}{6}$$

$$= \frac{2 \pm 2\sqrt{73}}{6}$$

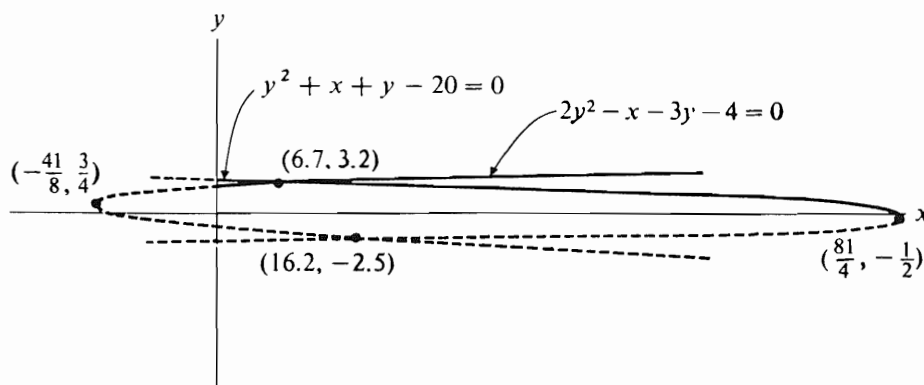
$$= \frac{1 \pm \sqrt{73}}{3}$$

$$x = -y^2 - y + 20$$

$$= -\frac{1 \pm 2\sqrt{73} + 73}{9} - \frac{1 \pm \sqrt{73}}{3} + 20$$

$$= -\frac{77 \pm 5\sqrt{73}}{9} + 20$$

Figure 1.82



$\sqrt{73} \approx 8.544$, so the approximate solutions are $(6.7, 3.2)$ and $(16.2, -2.5)$, and the approximate equilibrium point is $(6.7, 3.2)$.

Plotting the curves to verify the result geometrically (see Figure 1.82),

$$\begin{aligned} y^2 + y + x - 20 &= 0 & 2y^2 - x - 3y - 4 &= 0 \\ y^2 + y + \frac{1}{4} &= -x + 20 + \frac{1}{4} & 2(y^2 - \frac{3}{2}y + \frac{9}{16}) &= x + 4 + \frac{9}{8} \\ (y + \frac{1}{2})^2 &= -(x - \frac{81}{4}) & (y - \frac{3}{4})^2 &= \frac{1}{2}(x + \frac{41}{8}) \end{aligned}$$

EXAMPLE

Find the market equilibrium price and quantity for the following demand and supply equations (where x represents quantity and y represents price):

$$(x + 12)(y + 6) = 169$$

$$x - y + 6 = 0$$

Solving the equations simultaneously,

$$y = x + 6 \quad \text{from the second equation}$$

Substituting in the first equation,

$$(x + 12)(x + 12) = 169$$

$$x + 12 = \pm 13$$

$$x = 1, -25$$

$$y = x + 6$$

$$y = 7, -19$$

The solutions are $(1, 7)$ and $(-25, -19)$ and the equilibrium point is $(1, 7)$.

Plotting the curves to verify the result geometrically (see Figure 1.83),

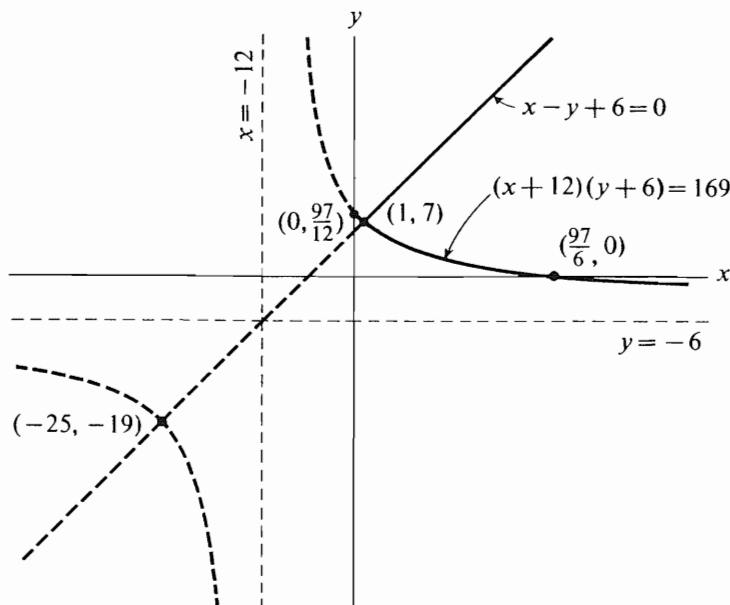
$$x - y + 6 = 0 \quad (x + 12)(y + 6) = 169$$

$$\text{If } x = 0, y = 6 \quad \text{asymptotes } x = -12, y = -6$$

$$\text{If } y = 0, x = -6 \quad \text{If } x = 0, y = \frac{97}{12}$$

$$\text{If } y = 0, x = \frac{97}{6}$$

Figure 1.83



PRODUCT TRANSFORMATION CURVES

Some production processes can yield more than one output. Sheep raising is a classic example of such a process: two outputs, wool and mutton, can be produced in varying proportions by a single production process. Many industrial production processes can also yield more than one output, for example, commodities that are similar but of different type or quality. A production or product transformation curve expresses the relationship between the quantities of two different commodities (joint products) produced by the same firm using common supplies of labor and raw materials. Note that the case of joint products is distinguished on a technical, not an organizational basis; the quantities of two or more joint products are technically or technologically interdependent. Cases in which a firm produces two or more technically independent products are inappropriate for analysis by product transformation curves.

A *product transformation curve* is defined as the locus of output quantity combinations which can be obtained from a given input. If the two output quantities produced are x and y , the product transformation curve relating them must be such that as one quantity increases the other quantity decreases. In addition, in order to fulfill certain reasonable economic assumptions, which need not be discussed here, product transformation curves are usually concave from below, as in Figure 1.84.

A product transformation curve is usually one of a family of possible product transformation curves, where the curves of the family correspond to various inputs. For example, Figure 1.84 shows product transformation curves represented by the first-quadrant parts of four of a family of concentric circles; the farther a curve lies from the origin, the greater the input to which it corresponds. Suitably located and oriented elliptic, parabolic, and hyperbolic curves are also appropriate in some cases for representing product transformation curves, as illustrated in Figures 1.85, 1.86,

Figure 1.84 Members of a family of circular product transformation curves.

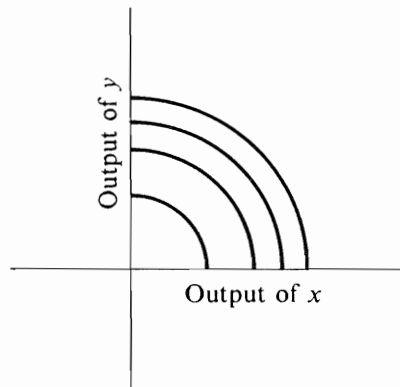


Figure 1.85 Example of an elliptic product transformation curve.

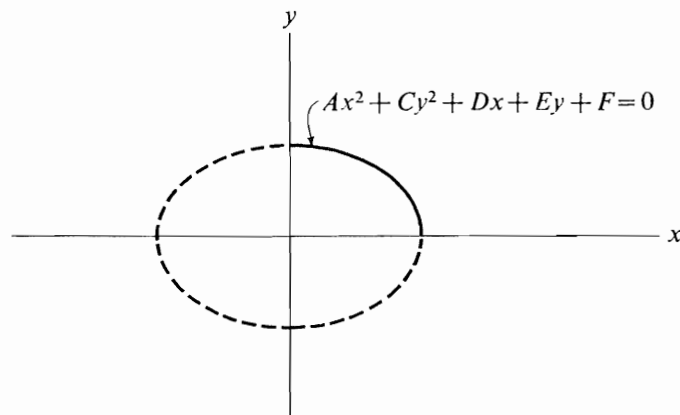


Figure 1.86 Examples of parabolic product transformation curves.

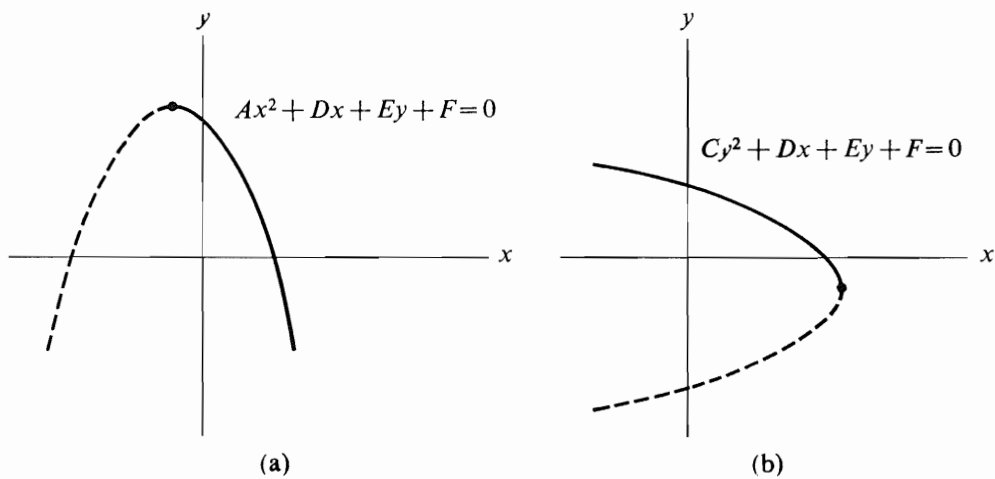
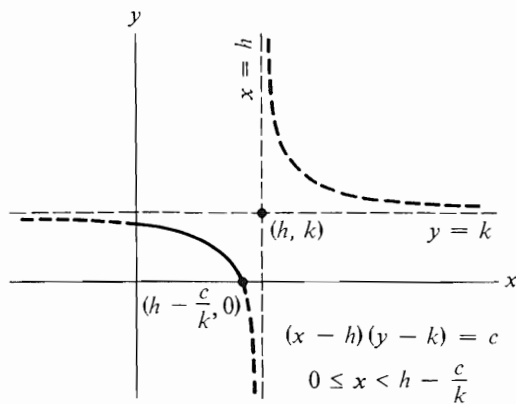


Figure 1.87 Example of a hyperbolic product transformations curve.



and 1.87. Note that for circular, elliptic, and parabolic product transformation curves, the entire first-quadrant part of the curve is appropriate. However, for hyperbolic product transformation curves, the range is limited to the lower branch of the hyperbola; this limitation is usually stated as a restriction on the range of x :

$$0 \leq x \leq h - \frac{c}{k}$$

where $(h - (c/k), 0)$ is the x -intercept.

EXAMPLE

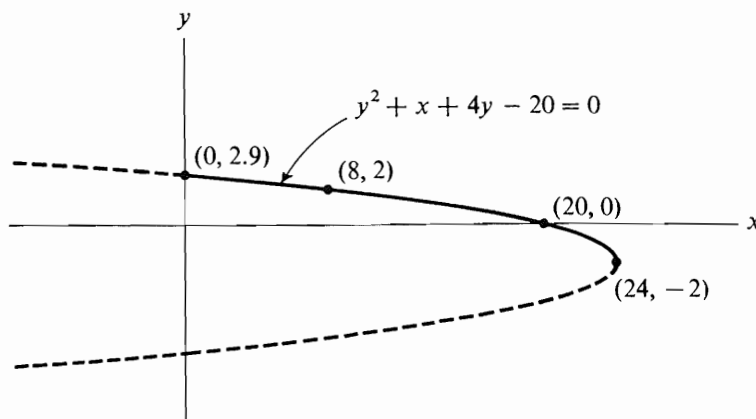
A company produces amounts x and y of two different grades of steel using the same production process. The product transformation curve for the input used is given by

$$y^2 + x + 4y - 20 = 0$$

(see Figure 1.88).

- What are the largest amounts x and y that can be produced?
- What amounts x and y should be produced in order to have $x = 4y$?

Figure 1.88



(a) x is as large as possible if $y = 0$, so the largest amount x is 20. y is as large as possible if $x = 0$, so the largest amount y is

$$\begin{aligned} y &= \frac{-4 \pm \sqrt{16 - (4)(-20)}}{2} \\ &= \frac{-4 \pm 4\sqrt{6}}{2} \\ &= -2 \pm 2\sqrt{6} \end{aligned}$$

$\sqrt{6} \approx 2.449$, so $y \approx 2.9$, -6.9 and the largest amount y is (approximately) 2.9.

(b) Substituting $x = 4y$ in $y^2 + x + 4y - 20 = 0$,

$$y^2 + 4y + 4y - 20 = 0$$

$$y^2 + 8y - 20 = 0$$

$$(y + 10)(y - 2) = 0$$

$$y = -10, 2$$

So the amounts produced are $x = 8$, $y = 2$.

EXAMPLE

A company produces amounts x and y of two different kinds of candy using the same production process. The product transformation curve for the input used is given by

$$5x^2 + 2y^2 = 98$$

(see Figure 1.89).

(a) What are the largest amounts x and y that can be produced?

(b) What amounts x and y should be produced in order to have $y = \frac{3}{4}x$?

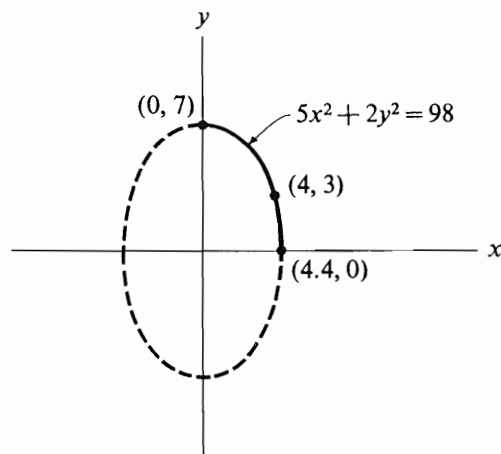
(a) If $y = 0$,

$$5x^2 = 98$$

$$x^2 = \frac{98}{5}$$

$$x = \pm \sqrt{\frac{98}{5}} = \pm 7\sqrt{\frac{2}{5}} = \pm \frac{7}{5}\sqrt{10}$$

Figure 1.89



$\sqrt{10} \approx 3.162$, so $x \approx \pm 4.4$ and the largest amount x is (approximately) 4.4. If $x = 0$,

$$2y^2 = 98$$

$$y^2 = 49$$

$$y = \pm 7$$

So the largest amount y is 7.

(b) Substituting $y = \frac{3}{4}x$ in $5x^2 + 2y^2 = 98$,

$$5x^2 + 2\left(\frac{9}{16}\right)x^2 = 98$$

$$\frac{49}{8}x^2 = 98$$

$$x^2 = 16$$

$$x = \pm 4, y = \pm 3$$

So the amounts produced are $x = 4, y = 3$.

EXAMPLE

A company produces amounts x and y of two petrochemicals using the same production process. The product transformation curve for the input used is given by

$$(x - 24)(y - 36) = 240 \quad x < \frac{52}{3}$$

(see Figure 1.90).

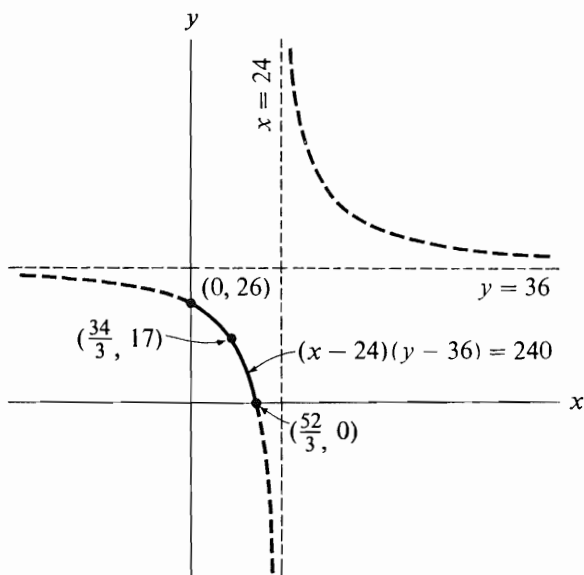
(a) What are the largest amounts x and y that can be produced?

(b) What amounts x and y should be produced to have $x = \frac{2}{3}y$?

(a) If $y = 0$,

$$x = -\frac{20}{3} + 24 = \frac{52}{3}$$

Figure 1.90



So the largest amount x is $\frac{52}{3}$. If $x = 0$,

$$y = -10 + 36 = 26$$

So the largest amount y is 26.

(b) Substituting $x = \frac{2}{3}y$ in $(x - 24)(y - 36) = 240$,

$$\left(\frac{2y}{3} - 24\right)(y - 36) = 240$$

$$(y - 36)(y - 36) = 360$$

$$y - 36 = \pm\sqrt{360}$$

$$y = 36 \pm 6\sqrt{10}$$

$$y \approx 36 \pm 19.0$$

$$y \approx 55, 17$$

So the amounts produced are (approximately) $x = \frac{34}{3}$, $y = 17$. (Note that the solution $x = \frac{110}{3}$, $y = 55$ violates the restriction $x \leq \frac{52}{3}$ and is on the upper branch of the hyperbola.)

EXAMPLE

A company produces amounts x and y of two different textiles using the same production process. The product transformation curve for the input used is given by

$$y = 20 - \frac{x^2}{5}$$

(see Figure 1.91).

- (a) What are the largest amounts x and y that can be produced?
 (b) What amounts x and y should be produced to have $x = y$?

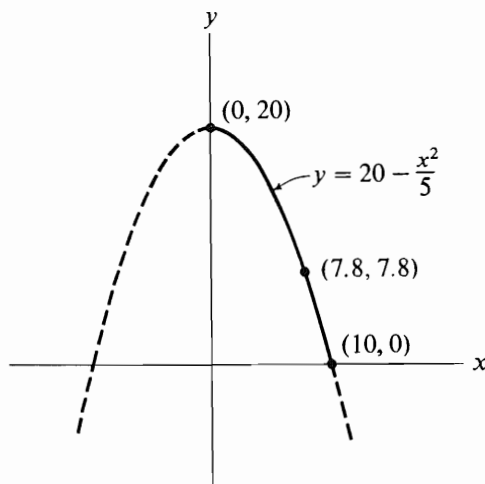
(a) If $y = 0$,

$$x^2 = 100$$

$$x = \pm 10$$

So the largest amount x is 10. If $x = 0$, $y = 20$, so the largest amount y is 20.

Figure 1.91



(b) Substituting $x = y$ in $y = 20 - \frac{x^2}{5}$,

$$5y = 100 - y^2$$

$$y^2 + 5y - 100 = 0$$

$$y = \frac{-5 \pm \sqrt{25 - (4)(-100)}}{2}$$

$$= \frac{-5 \pm 5\sqrt{17}}{2}$$

$$y \approx \frac{-5 \pm 5(4.123)}{2}$$

$$y \approx 7.8, -12.8$$

So the amounts produced are (approximately) $x = y = 7.8$.

PROBLEMS

For each of the following pairs of equations, i. determine which equation represents a demand curve and which equation represents a supply curve, ii. determine the market equilibrium quantity and price algebraically, and iii. check geometrically the algebraically determined equilibrium points.

1. a. $x = 16 - 2y$

b. $4x = 4y + y^2$

2. a. $x = 130 - 4y$

b. $y = 10 + \frac{x}{5} + \frac{x^2}{100}$

3. a. $y = 2 + \frac{x}{5} + \frac{x^2}{20}$

b. $y = \frac{30 - x}{4}$

4. a. $y = 16 - x^2$

b. $y = 4 + x$

5. a. $x = 32 - 4y - y^2$

b. $y = \frac{x}{20} + 1$

6. a. $y = 9x + 12$

b. $y = 39 - 3x^2$

7. a. $y = 6 + \frac{x^2}{4}$

b. $x = \sqrt{36 - y}$

8. a. $y = (x + 2)^2$

b. $y = 39 - 3x^2$

9. a. $y = 48 - 3x^2$

b. $y = x^2 + 4x + 16$

10. a. $x = 84 - y^2$

b. $x = y + 4y^2$

11. a. $x = 10y + 5y^2$

b. $x = 64 - 8y - 2y^2$

12. a. $x = 10y + 4y^2$

b. $x = 96 - 8y - 2y^2$

13. a. $(x + 16)(y + 12) = 480$

b. $y = 2x + 4$

14. a. $x = 2y^2 - 2y - 6$

b. $x = -y^2 - y + 18$

15. a. $y = 10 - 3x^2$

b. $y = 4 + x^2 + 2x$

16. a. $xy = 30$

b. $3y - x = 9$

17. a. $xy = 15$

b. $y = x + 2$

18. a. $(x + 10)(y + 20) = 300$

b. $x = 2y - 8$

19. a. $(x + 6)(y + 12) = 144$

b. $y = 2 + \frac{x}{2}$

20. a. $(x + 12)(y + 6) = 169$
b. $x - y + 6 = 0$
21. a. $(x + 5)(y + 6) = 80$
b. $y = \frac{x}{3} + 3$
22. a. $(x + 1)y = 5$
b. $y = \frac{x}{4}$
23. a. $x(y + 6) = 24$
b. $y - 2x + 4 = 0$
24. a. $y(x + 3) = 18$
b. $y - 3x + 6 = 0$
25. a. $(x + 4)(y + 2) = 24$
b. $y = 1 + \frac{x}{2}$
26. a. $y = x^2 + 5x + 1$
b. $y + 2x^2 - 9 = 0$
27. a. $x = 3y^2 - 3y - 2$
b. $x = 10 - y^2 - y$
28. a. $(x + 10)(y + 5) = 225$
b. $x - y + 5 = 0$

Each of the following equations represents a product transformation curve for the amounts x and y , respectively, of two related commodities; find the largest amounts x and y that can be produced.

29. $x = 36 - 6y^2$
30. $y = 65 - 12x - 5x^2$
31. $y = 45 - 9x^2$
32. $x = 16 - 4y - 2y^2$
33. The production manager has decided that the marketing department can sell 126 units of product daily, and he wants to produce that much. If he assumes that all factors other than the number of employees and the resultant output remain constant within the range of this total production, the production function can be expressed by the equation

$$2x^2 + 4x - y = 0$$

with x representing the number of employees and y representing the units of production. The production manager claims that he will need 7 men to produce the 126 units. a. Assuming that the equation is appropriate, is the production manager correct in his statement concerning the number of men needed? b. What type of curve does the equation represent? Graph the curve. c. Draw up a schedule of units of production output per man employed in the range 1 through 7 employees. Indicate the change in the number of units produced through this range as each employee is added.

34. A company's operations research director believes that the short-run average cost of production can be expressed by the equation

$$x^2 - 16x - y + 68 = 0$$

with x representing the number of units produced and y representing the average cost per unit. He says that the average cost will be lowest when eight units are produced. a. Is his statement correct? b. What type of curve is represented? Graph the curve. c. Draw up a schedule of values of y for the range $x = 4$ through $x = 12$ and indicate the amount of change in y for each change in x .

35. In national income analysis, the demand for a stock of money to be held, or the "liquidity preference," as Keynes calls it, is often considered to be dependent on three motives: the transactions motive, the precautionary motive, and the speculative motive. Assume that, at a given level of national income, the effects of the transactions and precautionary motives are constant. The speculative motive is considered to be a function of the interest rate as expressed by the equation $(x - 1)y = 4$, where x is the rate of interest (%) and y is the demand for money to hold expressed in billions of dollars. a. What type of curve does the equation express? Graph the curve. b. Draw up a schedule of values for y , the amount of money demanded to be held (in billions of dollars) for the values of x from 2% through

- 7%. What is the value of y when $x = 100$ (in billions of dollars)? c. Note and describe the segment of the curve which represents the "liquidity trap"—that segment in which the interest rate seems to lose its force as an effective factor in influencing the demand for money to hold.
36. According to convention, in the economic analysis discussed in Problem 35, the dependent variable (the demand for money to be held) is often assigned to the x -axis rather than the y -axis. An equation used to express Keynesian ideas concerning the relationship between the rate of interest and the money demanded to be hld is then $x(y - 1) = 4$. a. What type of curve does this equation represent? Graph the curve. b. Draw up a schedule of values for the interest rate y for the values of x from 1 through 7 (in billions of dollars). What is the value of y when $x = 100$ (in billions of dollars)? c. Note and describe the segment of the curve that represents the "liquidity trap."
37. Consider the parabola $y = \frac{1}{4}x^2 - x + 4$ and the equilateral hyperbola $(x + 2)(y - 2) = 4$. a. Show that for $x = 0$ and $x = 2$ both equations have the same value of y , but for $x = 4$ and $x = -2$ they have different values of y . b. Show that the two equations have the same value of y only for $x = 0$ and $x = 2$. c. Sketch the two curves on the same set of axes.
38. A company produces x and y amounts of steel of two different grades, using the same resources; the product transformation curve is
- $$y = 20 - \frac{300}{30 - x} \quad (x < 30)$$
- a. Sketch the curve. b. Determine the largest amounts of x and y that can be produced. c. If the demand for grade x steel is twice that for grade y steel, determine the amounts the company should produce.
39. A company manufactures two grades of candy from the same resources. If x and y represent the quantities produced, the product transformation curve is
- $$(x - 24)(y - 36) = 240 \quad (x < 24)$$
- a. Sketch the curve. b. Determine the largest amounts of x and y that can be produced. c. If the demand for grade x candy is two-thirds that for grade y , determine the amounts the company should produce.
40. A company manufactures two grades of paper from the same resources; if x and y represent the quantities produced, the product transformation curve is
- $$(x - 30)(y - 15) = 150 \quad (x < 30)$$
- a. Sketch the curve. b. If the demand for grade x paper is three times that for grade y paper, determine the amounts of paper the company should produce. c. If the demand for grade y paper exceeds that for grade x paper by four units, determine the amounts the company should produce.

Answers to Odd-Numbered Problems

- | | |
|---------------|---------------|
| 1. a. demand | 11. a. supply |
| b. supply | b. demand |
| (8, 4) | (40, 2) |
| 3. a. supply | 13. a. demand |
| b. demand | b. supply |
| (6.91, 5.77) | (4, 12) |
| 5. a. demand | 15. a. demand |
| b. supply | b. supply |
| (20, 2) | (1, 7) |
| 7. a. supply | 17. a. demand |
| b. demand | b. supply |
| (4.90, 12) | (3, 5) |
| 9. a. demand | 19. a. demand |
| b. supply | b. supply |
| (2.37, 31.15) | (3.22, 3.61) |

21. a. demand
b. supply
(3, 4)
23. a. demand
b. supply
(3, 2)
25. a. demand
b. supply
(2, 2)
33. a. yes
b. parabola, vertex $(-1, -2)$, axis parallel to y -axis

27. a. supply
b. demand
(4, 2)
29. $x_{\max} = 36$
 $y_{\max} = \sqrt{6}$
31. $x_{\max} = \sqrt{5}$
 $y_{\max} = 45$

Number of employees x	Total units produced y	Δy
1	6	
2	16	10
3	30	14
4	48	18
5	70	22
6	96	26
7	126	30

35. a. equilateral hyperbola, transverse axis parallel to y -axis
center $(1, 0)$
lies in first and third quadrants of its asymptotes
asymptotes: $x = 1, y = 0$
intercept: $(0, -4)$

Interest rate x	Amount of money demanded to be held y
2	4
3	2
4	$\frac{4}{3}$
5	1
6	$\frac{4}{3}$
7	$\frac{2}{3}$
100	$\frac{4}{99}$

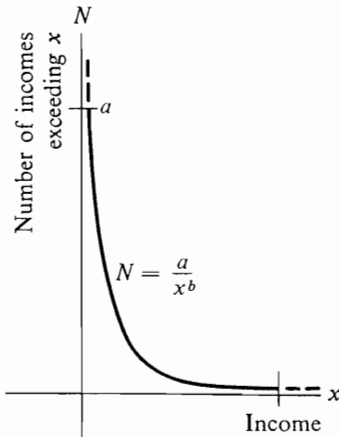
- c. segment for which $x > 2$
39. b. $x_{\max} = 17\frac{1}{3}, y_{\max} = 26$
c. $x = 11.36, y = 17.04$

PARETO'S LAW OF DISTRIBUTION OF INCOME

The economist Vilfredo Pareto proposed the following law of distribution of income:
The number of individuals N from a given population of size a whose income exceeds x is

$$N = \frac{a}{x^b}$$

Figure 1.92



where b is a population parameter, usually approximately 1.5. This equation represents a generalized equilateral hyperbola and is appropriate only for the range $0 < N \leq a$ and $0 < x < \text{maximum income in the population}$ (see Figure 1.92).

Note that Pareto's equation implies that both N and x are continuous; in fact, although x might be considered approximately continuous, N is by definition discrete. For convenience, relationships involving discrete variables are frequently represented as continuous; the discrepancy between the discrete nature of a variable and its continuous representation must be considered in interpreting analyses based on such representations.

For above-subsistence incomes, data indicate that Pareto's law is generally fairly accurate. Pareto suggested that the value of b is approximately 1.5; data indicate that this varies from population to population, but 1.5 is usually a good approximation.

NOTE: The notation used in the following examples is reviewed in the next section, which deals with exponential and logarithmic functions. Note, however, that Pareto's law is represented by a power function, not by an exponential function.

EXAMPLE

Pareto's law for the distribution of incomes for a particular group is

$$N = \frac{216 \times 10^{10}}{x^{3/2}}$$

- How many people are millionaires?
- How many people have incomes between \$3600 and \$10,000?
- What is the lowest income of the 80 people with the highest incomes?

$$(a) \ N = \frac{216 \times 10^{10}}{(10^6)^{3/2}} = 2160 \text{ millionaires}$$

- The number with incomes exceeding \$3600 is

$$N = \frac{216 \times 10^{10}}{(3600)^{3/2}} = \frac{216 \times 10^{10}}{6^3 \times 10^3} = 10^7$$

The number with incomes exceeding \$10,000 is

$$N = \frac{216 \times 10^{10}}{(10^4)^{3/2}} = 216 \times 10^4$$

so the number with incomes between \$3600 and \$10,000 is $10^7 - 216 \times 10^4 = (1000 - 216) \times 10^4 = 784 \times 10^4$ or 7,840,000.

$$(c) \quad 80 = \frac{216 \times 10^{10}}{x^{3/2}}$$

$$x^{3/2} = \frac{216 \times 10^{10}}{80} = 27 \times 10^9$$

$$x = (27 \times 10^9)^{2/3} = 9 \times 10^6$$

or \$9,000,000 is the lowest income of the 80 people with the highest incomes.

EXAMPLE

Pareto's law for the distribution of incomes for a given population is

$$N = \frac{16 \times 10^{12}}{x^{5/3}}$$

- (a) How many people have incomes below \$8000?
- (b) How many people have incomes over \$125,000 but less than \$1,000,000?
- (c) What is the lowest income of the 50 people with the highest incomes?

$$(a) \quad N = \frac{16 \times 10^{12}}{(8 \times 10^3)^{5/3}} = \frac{16 \times 10^{12}}{2^5 \times 10^5} = 5 \times 10^6$$

so 5×10^6 people have incomes above \$8000 and $16 \times 10^{12} - (5 \times 10^6) = (16 \times 10^6 - 5) \times 10^6$ or 15,999,995,000,000 have incomes below \$8000.

- (b) The number with incomes over \$125,000 is

$$\begin{aligned} N &= \frac{16 \times 10^{12}}{(5^3 \times 10^3)^{5/3}} = \frac{16 \times 10^{12}}{5^5 \times 10^5} = \frac{16 \times 10^7}{5^5} \\ &= 16 \times 2^5 \times 10^2 = 512 \times 10^2 \text{ or } 51,200 \end{aligned}$$

The number with incomes over \$1,000,000 is

$$N = \frac{16 \times 10^{12}}{(10^6)^{5/3}} = 16 \times 10^2 \text{ or } 1600$$

So the number with incomes over \$125,000 but less than \$1,000,000 is $51,200 - 1600 = 49,600$.

$$(c) \quad 50 = \frac{16 \times 10^{12}}{x^{5/3}}$$

$$x^{5/3} = \frac{16 \times 10^{12}}{50} = 32 \times 10^{10}$$

$$x = (32 \times 10^{10})^{3/5} = 8 \times 10^6$$

or \$8,000,000 is the lowest income of the 50 people with the highest incomes.

PROBLEMS

- 1. Pareto's law for the distribution of incomes for a particular group is

$$N = \frac{8 \times 10^8}{x^{3/2}}$$

- a. How many have incomes exceeding \$1600?
- b. How many have incomes between \$1600 and \$3600?
- c. What is the lowest income of the 800 who have the highest incomes?

2. Pareto's law for the distribution of incomes for a particular group is

$$N = \frac{1.9 \times 10^{12}}{x^{1.70}}$$

- How many have incomes exceeding \$50,000?
 - How many have incomes between \$25,000 and \$50,000?
 - What is the lowest income of the million who have the highest incomes?
3. Pareto's law for the distribution of incomes for a particular group is

$$N = \frac{100,000}{x^2}$$

- How many have incomes exceeding \$15?
 - How many have incomes between \$50 and \$75?
 - What is the lowest income of the 5 who have the highest incomes?
4. Pareto's law for the distribution of incomes for a particular group is

$$N = \frac{32 \times 10^{10}}{x^{4/3}}$$

- How many have incomes between \$125,000 and \$1,000,000?
 - What is the lowest income of the 200 who have the highest incomes?
5. Pareto's law for the distribution of incomes for a particular group is

$$N = \frac{625 \times 10^9}{x^{3/2}}$$

- How many have incomes between \$2500 and \$10,000?
 - What is the lowest income of the 5000 who have the highest incomes?
6. Pareto's law for the distribution of incomes for a particular group is

$$N = \frac{6 \times 10^9}{x^{3/2}}$$

- How many have incomes exceeding \$2500?
 - How many have incomes between \$2500 and \$10,000?
 - What is the lowest income of the 6 who have the highest incomes?
7. Pareto's law for the distribution of incomes is

$$N = \frac{a}{x^b}$$

- How many have incomes exceeding k ?
 - How many have incomes exceeding \$100,000?
 - What is the lowest income of the 10 who have the highest incomes?
 - How many have incomes between s and t ?
 - How many have incomes between \$500,000 and \$1,500,000?
8. If the demand function is

$$y = \frac{17.6}{x^{1.43}}$$

(where x is quantity demanded and y is price):

- Find the price if the quantity demanded is 2.
 - Find the quantity demanded if the price is \$5.
9. If the demand function is

$$x = \frac{12.03}{y^{0.21}}$$

(where x is quantity demanded and y is price):

Find the price if the quantity demanded is a. 6, b. 3.3.

Find the quantity demanded if the price is c. \$100, d. \$67.

10. If the demand function is

$$x = \frac{a}{y^b}$$

(where x is quantity demanded and y is price):

Find the price if the quantity demanded is a. $5a$, b. 1.

Find the quantity demanded if the price is c. \$1, d. a .

Answers to Odd-Numbered Problems

1. a. 12,500 c. $\left(\frac{a}{10}\right)^{1/b}$
 b. 8796 d. $\frac{a(t^b - s^b)}{s^b t^b}$
 c. \$10,000 e. $\frac{a(3^b - 1)}{(15 \times 10^5)^b}$
3. a. 444 9. a. \$27.455
 b. 22 b. \$473.21
 c. \$141 c. 4.574
5. a. 4,375,000 d. 4.975
 b. \$250,000
7. a. $\frac{a}{k^b}$
 b. $\frac{a}{10^{5b}}$

1.8 Exponential and Logarithmic Curves

Exponential and logarithmic curves are closely related. Graphical representations of exponential and logarithmic functions are discussed in this section.

EXPONENTIAL FUNCTIONS

A function having a variable base and a constant exponent is a *power function*. For example, $y = x^a$ is a power function, where x is the base and a is the exponent. Many examples of power functions have been discussed in previous sections. A function having a constant base and a variable exponent is an *exponential function*. For example, $y = a^x$ is an exponential function, where a is the base and x is the exponent.

The following rules of exponents are applicable to both power functions and exponential functions.

If x , y , and z are positive real numbers,

Rule

1. $x^m x^n = x^{m+n}$
2. $(x^m)^n = x^{mn}$
3. $(xyz)^m = x^m y^m z^m$
4. $\left(\frac{x}{y}\right)^m = \frac{x^m}{y^m}$

Example

$$10^2 \times 10^3 = 10^5$$

$$(2^4)^6 = 2^{24}$$

$$(2 \times 4 \times 5)^6 = 2^6 \times 4^6 \times 5^6$$

$$\left(\frac{3}{8}\right)^4 = \frac{3^4}{8^4}$$

$$5. \ x^{-m} = \frac{1}{x^m}$$

$$5^{-6} = \frac{1}{5^6}$$

$$6. \ \frac{x^m}{x^n} = x^{m-n} = \frac{1}{x^{n-m}}$$

$$\frac{4^5}{4^8} = 4^{-3} = \frac{1}{4^3}$$

$$7. \ x^{1/n} = \sqrt[n]{x}, \text{ where } n \text{ is a positive integer, read the } n\text{th root of } x$$

$$7^{1/3} = \sqrt[3]{7}$$

$$8. \ x^{m/n} = \sqrt[n]{x^m} = (\sqrt[n]{x})^m, \text{ where } m \text{ and } n \text{ are integers, } n > 0, \text{ read the } m\text{th power of the } n\text{th root of } x.$$

$$9^{5/8} = \sqrt[8]{9^5} = (\sqrt[8]{9})^5$$

NOTE: By definition, if $x \neq 0$, $x^0 = 1$; that is, any base to the zero power is 1.

One very convenient use of exponents is in the representation of numbers which are very large or very small in absolute value, by expressing them as a constant multiplied by an exponent of 10. For example, 80,000 can be written 8×10^4 ; 5,000,000 can be written 5×10^6 . Similarly, the fraction 0.002 can be written 2×10^{-3} ; the fraction 0.0000000096 can be written 9.6×10^{-10} . This representation of numbers is referred to as *scientific notation*; the significant digits in the number are conventionally expressed with one digit to the left of the decimal point. This notation was used in the examples of Pareto's law, but the convention concerning decimal points was violated for convenience of calculation.

The simplest exponential functions are of the form $y = b^x$, $b > 0$. The curve representing the function $y = b^x$ lies entirely in the first two quadrants; it is monotonically decreasing if $0 < b < 1$ and monotonically increasing if $b > 1$; in both cases the curve is asymptotic to the x -axis and has y -intercept $(0, 1)$. The parameter b determines the curvature of the function (see Figure 1.93). Note that if $b = 1$, $y = b^x$ becomes $y = 1$, a straight line parallel to the x -axis.

The most frequently used exponential function is $y = e^x$, where e is the base of the natural logarithms, approximately 2.718. Almost all the exponential functions used in economic theory have the base e . The precise definition of this base and its advantages are discussed in connection with logarithms in the following sections.

Figure 1.93

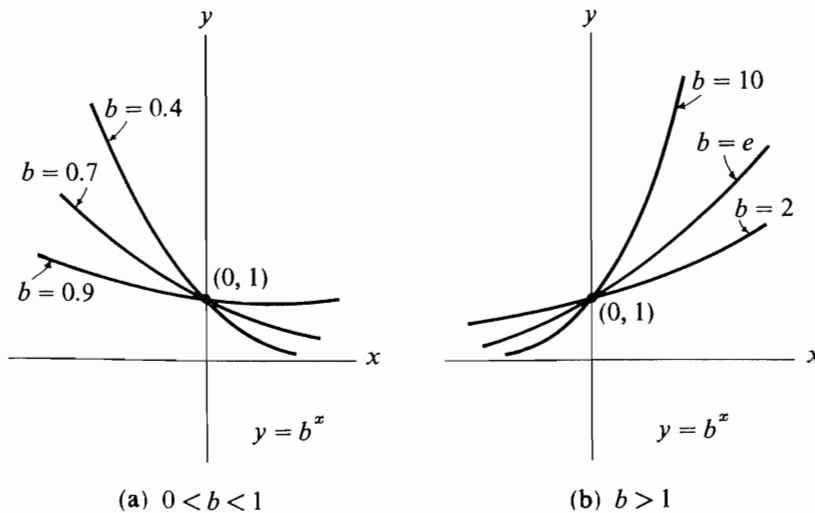
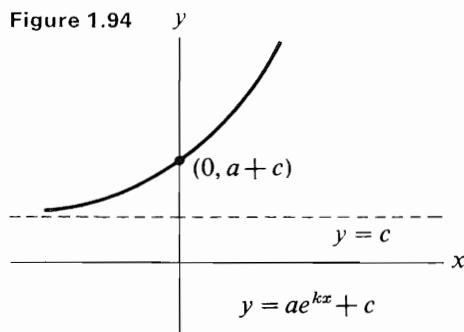
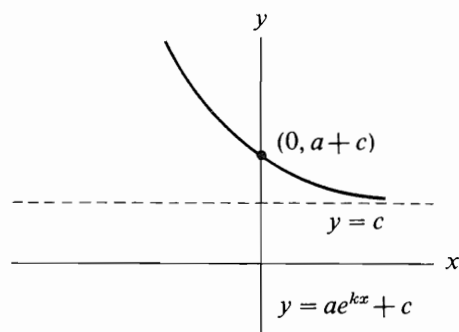
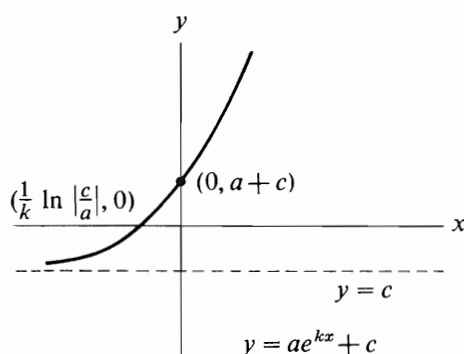
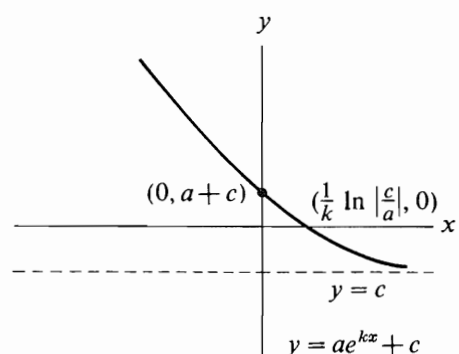
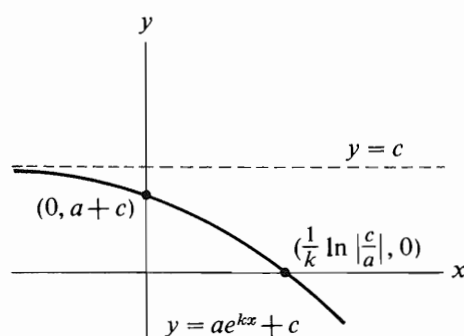
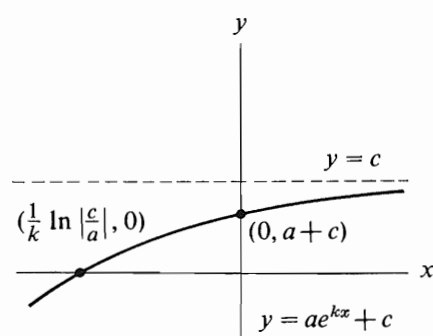
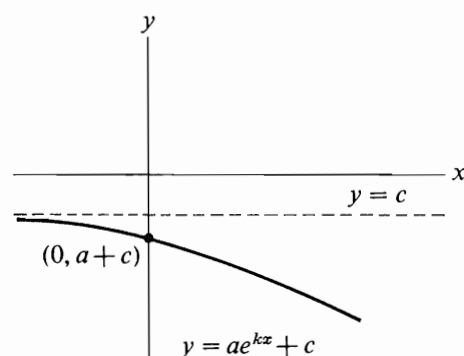
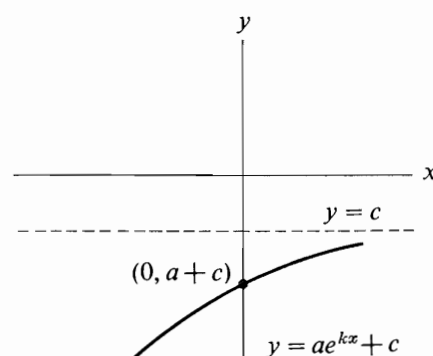


Figure 1.94

(a) $a > 0, k > 0, c \geq 0$ (b) $a > 0, k < 0, c \geq 0$ (c) $a > 0, k > 0, c \leq 0, |c| < a$
(if $|c| > a$, intercepts negative)(d) $a > 0, k < 0, c \leq 0, |c| < a$
(if $|c| > a$, intercepts negative)(e) $a < 0, k > 0, c > 0, c > |a|$
(if $c < |a|$, intercepts negative)(f) $a < 0, k < 0, c > 0, c > |a|$
(if $c < |a|$, intercepts negative)(g) $a < 0, k > 0, c \leq 0$ (h) $a < 0, k < 0, c \leq 0$

A more general form for exponential functions is

$$y = ae^{kx} + c$$

The curve representing the function $y = ae^{kx} + c$ is asymptotic to the line $y = c$; it approaches its asymptote from above if $a > 0$ and from below if $a < 0$. The curve is monotonically increasing if a and k are of the same sign and monotonically decreasing if a and k are of opposite signs. The curve has y -intercept $(0, a + c)$ and, if $c < 0$, x -intercept $((1/k) \ln |c/a|, 0)$; see Figure 1.94. (The x -intercept is obtained in Technical Note V.)

LOGARITHMIC FUNCTIONS

In 1614, John Napier wrote the first treatise on logarithms. He based his discussion on the comparison of two sets of numbers—one in arithmetic progression and one in geometric progression. For example, consider the following two progressions:

Arithmetic: 0 1 2 3 4 5 6 ...
Geometric: 1 3 9 27 81 243 729 ...

If the numbers of the arithmetic progression are considered as exponents (powers) of 3, the corresponding numbers of the geometric progression result from raising 3 to the indicated power. Thus

$$3^0 = 1$$

$$3^1 = 3$$

$$3^2 = 9$$

$$3^3 = 27$$

and so forth. Furthermore, the two progressions indicate that in order to multiply, it is necessary only to add exponents: $3^2 \times 3^3 = 3^{2+3} = 3^5$. And to divide, it is necessary only to subtract exponents: $3^5 \div 3^2 = 3^{5-2} = 3^3$. This idea is the basis for the computational short cuts possible using logarithms.

By definition, the *logarithm* of a positive number y to a positive base b , other than 1, is the exponent x to which the base must be raised to obtain the number. Thus, if $b > 0$, $b \neq 1$, and b , x , and y are related by the equation

$$y = b^x \tag{3}$$

then x , the exponent of b , is the logarithm of y to the base b . This relationship may also be expressed by the equation

$$x = \log_b y \tag{4}$$

Equations (3) and (4) express the same relationship between b , x , and y in exponential and logarithmic forms, respectively, and are inverse functions—that is, if y is an exponential function of x , then x is a logarithmic function of y .

Although the base of a logarithm can be any positive number other than 1, in practice the base is almost always either 10 (common or Briggsian logarithms) or $e \approx 2.718$ (natural or Napierian logarithms). The base e is defined in terms of limits as follows:

$$e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n} \right)^n$$

The concept of limits is discussed in detail later; for present purposes it is sufficient to

note that e can be approximated as accurately as desired by increasing n in the expression $(1 + 1/n)^n$.

In some cases, common logarithms are more convenient for computational work; natural logarithms are more convenient for theoretical work. By convention, $\log x$ denotes the common logarithm of x and $\ln x$ denotes the natural logarithm of x . If any other base is meant, it is specified.

The use of logarithms can save considerable work in computations involving numbers that are very large or very small in absolute value. The essential properties of logarithms are summarized in the following rules:

If x and y are positive real numbers,

Rule	Example
1. $\log_b(xy) = \log_b x + \log_b y$	$\log_4 10 = \log_4 5 + \log_4 2$
2. $\log_b\left(\frac{x}{y}\right) = \log_b x - \log_b y$	$\log_8\left(\frac{5}{6}\right) = \log_8 5 - \log_8 6$
3. $\log_b x^n = n \log_b x$	$\log_5 3^2 = 2 \log_5 3$
4. $\log_b \sqrt[n]{x} = \frac{1}{n} \log_b x$	$\log_9 \sqrt[6]{2} = \frac{1}{6} \log_9 2$
5. $\log_a x = \log_a b (\log_b x)$ $= \left(\frac{1}{\log_b a}\right) \log_b x$	$\log_{10} 14 = (\log_{10} 2)(\log_2 14)$ $= \left(\frac{1}{\log_2 10}\right) \log_2 14$

The simplest logarithmic functions, mentioned above, are of the form

$$y = \log_b x, \quad b > 0 \text{ and } b \neq 1$$

The curve representing the function $y = \log_b x$ lies entirely in the first and fourth quadrants; it is monotonically decreasing if $0 < b < 1$ and monotonically increasing if $b > 1$; in both cases the curve is asymptotic to the y -axis and has x -intercept $(1, 0)$. The parameter b determines the curvature of the function (see Figure 1.95).

Figure 1.95

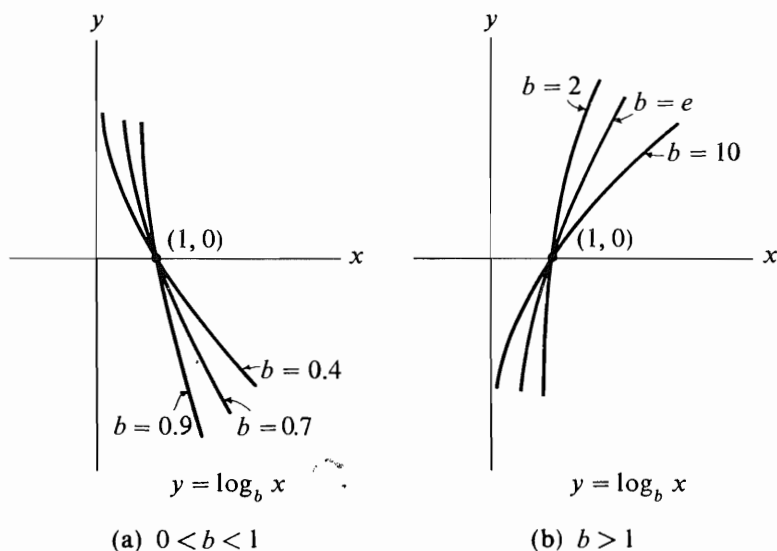


Figure 1.96

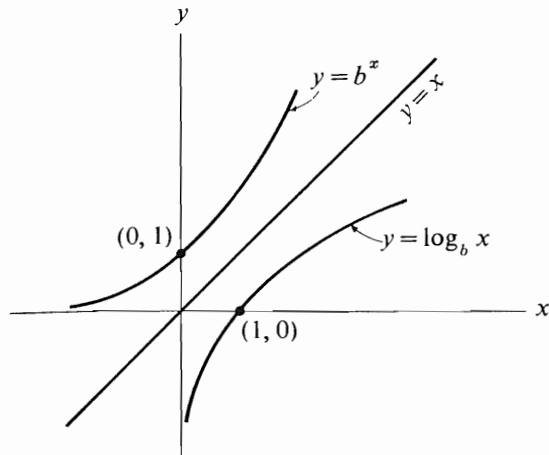
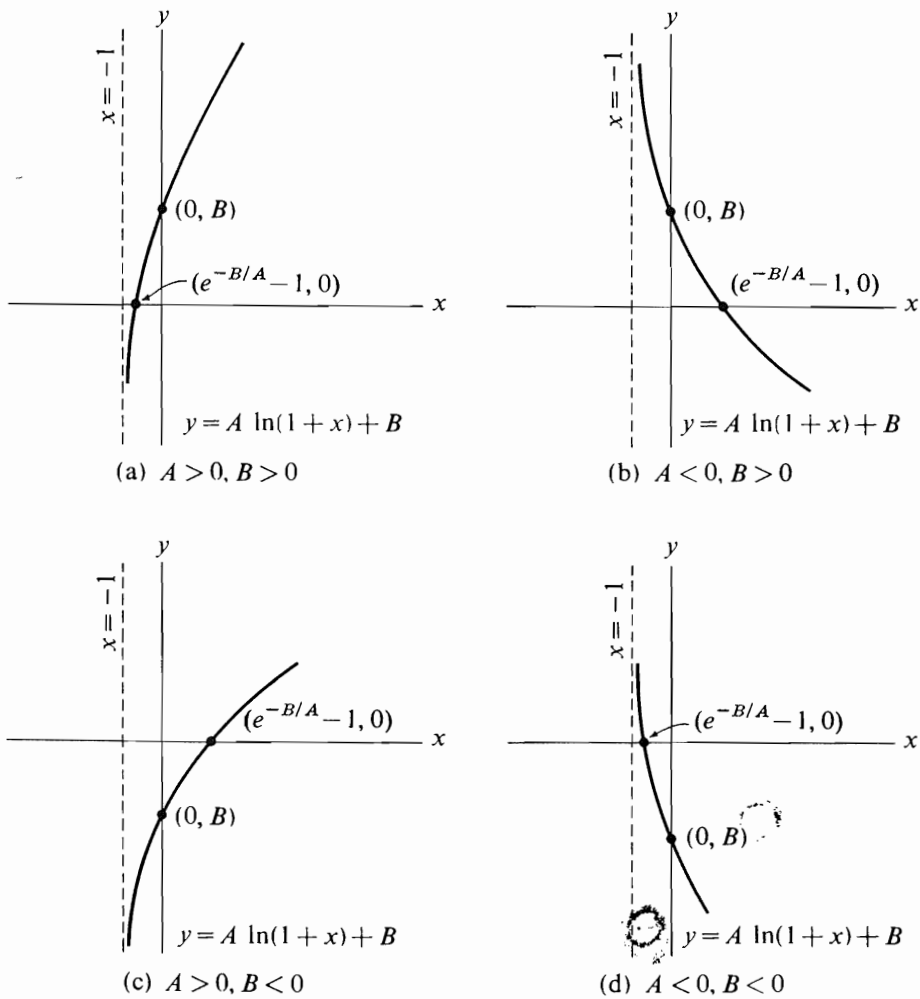


Figure 1.97



Since $y = b^x$ and $y = \log_b x$ are inverse functions, the graph of either one of these families of curves (or of one particular curve for some specified value of b) can be obtained from the graph of the other geometrically by reflection about the line $y = x$. (Recall that reflection of a curve about a line consists of obtaining the curve that is symmetric to the given curve with respect to the line; see Figure 1.96.)

A more general form for logarithmic functions is

$$y = A \ln(1 + x) + B \quad x > -1$$

The curve representing the function $y = A \ln(1 + x) + B$ lies to the right of and is asymptotic to the line $x = -1$; the curve is monotonically increasing if $A > 0$ and monotonically decreasing if $A < 0$. It has x -intercept $(e^{-B/A} - 1, 0)$ and y -intercept $(0, B)$; see Figure 1.97. (The x -intercept is obtained in Technical Note V.)

PROBLEMS

- For what ranges of x are the following curves defined?
 - $y = \ln(x + 1)$
 - $y = \ln(x - 4)$
 - $y = \ln(x^3 + 27)$
 - $y = \ln(x^4 - 16)$
- Sketch on the same set of axes the curves represented by the equations
 - $y = 2^x$
 - $y = 2^{-x}$
 - $y = 2^{x/2}$
 - $y = 2^{2x}$
- Sketch on the same set of axes the curves represented by $y = a^x$ for $a = 2, e, 3$, and 10 .
- Sketch on the same set of axes the curves represented by $y = \log_a x$ for $a = 2, e, 3$, and 10 .
- Sketch on the same set of axes the curves represented by the equations
 - $y = 3^x$
 - $y = \log_3 x$
- For what ranges of x are the following curves defined?
 - $y = \frac{1}{2} \ln(1 - x^2)$
 - $y = \ln(36 - x^2)$
 - $y = \ln(x^2 - 36)$
 - $y = -3 - 6 \ln(x - 2)$
 - $y = \ln(x - 27)$

Answers to Odd-Numbered Problems

- $x > -1$
 - $x > 4$
 - $x > -3$
 - $x < -2$ and $x > 2$

1.9 Applications of Exponential and Logarithmic Curves in Business and Economics

As can be noted in Figures 1.94 and 1.97, several types of exponential and logarithmic functions are appropriate for representing demand or supply curves. Exponential cost functions, representing total cost as an exponential function of number of units produced, are also frequently used. Many examples of these uses of exponential and logarithmic curves appear in later sections. The applications discussed in this section are compound interest and various types of growth functions, problems for which exponential and logarithmic functions are uniquely appropriate. A table of exponential functions and a table of logarithms are provided at the end of the book.

COMPOUND INTEREST

If the interest rate is $100i$ percent per year payable (that is, compounded) k times per year, an amount of money x (the principal) becomes after n years

$$y = x \left(1 + \frac{i}{k} \right)^{nk}$$

(when k is large, $y \approx xe^{in}$, where $e \approx 2.718$ is the base of the natural logarithms).

This formula can be developed as follows: If the interest at rate $100i$ percent is payable yearly, then the amount (principal plus interest) y_1 at the end of the first year is

$$y_1 = x + ix = x(1 + i)$$

The amount y_2 after 2 years is

$$y_2 = [x(1 + i)][1 + i] = x(1 + i)^2$$

The amount y_3 after 3 years is

$$y_3 = [x(1 + i)^2][1 + i] = x(1 + i)^3$$

\vdots

The amount y_n after n years is

$$y_n = x(1 + i)^n$$

If interest is payable k times per year, then the rate of interest in any one period is $100(i/k)$ percent and the number of periods is nk . Thus the amount after n years is

$$y = x \left(1 + \frac{i}{k} \right)^{nk}$$

EXAMPLE

A man deposits \$5000 at 4% interest. How much has he (principal plus interest) after 10 years

(a) if interest is payable yearly, and (b) if interest is payable quarterly?

$$(a) \quad y = x(1 + i)^n$$

$$= 5000(1 + 0.04)^{10}$$

$$\log y = \log 5000 + 10 \log 1.04$$

$$= 3.6990 + (10)(0.0170)$$

$$= 3.8690$$

$$y = \$7396.67$$

$$(b) \quad y = x \left(1 + \frac{i}{k} \right)^{nk}$$

$$= 5000 \left(1 + \frac{0.04}{4} \right)^{40}$$

$$\log y = \log 5000 + 40 \log 1.01$$

$$= 3.6990 + 40(0.0043)$$

$$= 3.8710$$

$$y = \$7430.00$$

GROWTH FUNCTIONS

There are many relationships in business and economics which are appropriately represented by curves referred to as growth functions, for example, number of employees as a function of annual sales of a company, amount of finished stock as a function of days after beginning a production run, sales as a function of advertising expenditure, maintenance cost as a function of number of hours a machine is run, and sales as a function of length of time a product has been on the market can be represented by growth functions.

The basic property of the various curves referred to as growth functions is that they are monotonically increasing. Growth functions may or may not have an upper asymptote (although usually for business and economic applications absence of an upper asymptote is reasonable only when relatively short periods of time are considered) and they may be of various shapes. Three particular types of growth functions are discussed below: (1) very simple functions originally developed to describe certain types of biological growth and appropriate only for growth that does not have an upper asymptote, (2) Gompertz functions used to describe growth that starts rather slowly and approaches an upper asymptote, and (3) learning functions originally used by psychologists to describe human learning and appropriate for growth that begins rapidly, levels off, and approaches an asymptote.

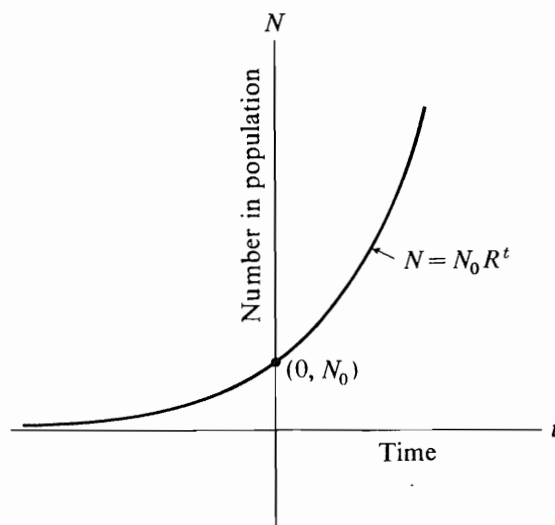
Biological Growth Curves

Many laws of biological growth are represented by the equation

$$N = N_0 R^t$$

where N is the number of individuals in the population at time t , N_0 is the initial number of individuals in the population (at time zero), and $R > 0$ is the rate of growth (see Figure 1.98). This equation is based on the model of a population each of whose members produces $R - 1$ additional members in each unit of time and none of whose members die. (This situation is approximated for certain laboratory cultures, at least for limited periods of time.) The equation can be developed as follows:

Figure 1.98



Time	Number in Population
0	N_0
1	$N_0 + N_0(R - 1) = N_0 R$
2	$N_0 R + N_0 R(R - 1) = N_0 R^2$
3	$N_0 R^2 + N_0 R^2(R - 1) = N_0 R^3$
\vdots	\vdots
t	$N_0 R^{t-1} + N_0 R^{t-1}(R - 1) = N_0 R^t$

Organization theorists have claimed that the function $N = N_0 R^t$ is appropriate for describing the early growth of a rapidly developing company; because it has no upper asymptote, the function is inappropriate for describing biological or organizational growth for an indefinite period of time. In addition, the behavior of the curve, as illustrated in the example below, indicates that it should be used with extreme caution in theoretical analyses.

EXAMPLE

The National Aerospace Research Organization is beginning its operations with a staff of 5 men. At the end of each year of its operation, each employee will hire 3 assistants. How many employees will National Aerospace Research Organization have after 10 years of operation?

$$\begin{aligned}
 N &= N_0 R^t \\
 &= (5)(4^{10}) \\
 &= 1,048,576 \text{ employees after 10 years}
 \end{aligned}$$

Gompertz Curves

Gompertz curves, named for their originator, are represented by the equation

$$N = ca^{R^t}$$

where N is the number of individuals in the population at time t , R ($0 < R < 1$) is the rate of growth, a is the proportion of initial growth, and c is the growth at maturity (that is, the upper asymptote). Note that when $t = 0$, $N = ca$, which corresponds to N_0 of the biological growth function.

The two basic types of Gompertz curves are characterized as follows and illustrated in Figure 1.99:

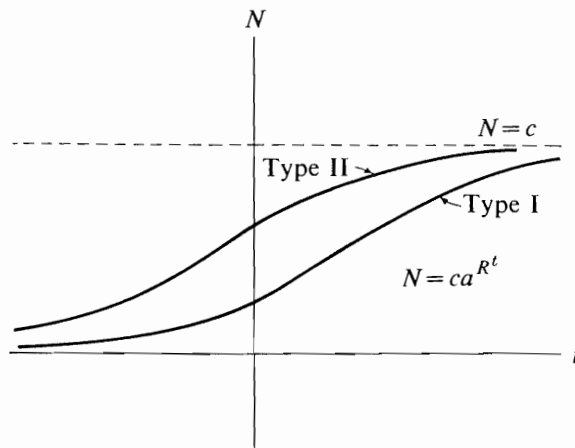
$$\text{Type I: } 0 < a < \frac{1}{e}$$

$$\text{Type II: } \frac{1}{e} \leq a < 1$$

Curves of Type I are positively accelerated for small positive t and negatively accelerated for large positive t ; curves of type II are negatively accelerated for all positive t . The mathematical definitions of positive and negative acceleration are discussed later; essentially, a positively accelerated curve is increasing at an increasing rate and a negatively accelerated curve is increasing at a decreasing rate.

Gompertz curves have been used extensively by psychologists to describe various aspects of human growth and development, including some types of learning. Organization theorists have found Gompertz curves appropriate for describing the growth

Figure 1.99



of many types of organizations. These curves are also appropriate for many other functions in business and economics, for example, total revenue and production functions.

EXAMPLE

On the basis of expected sales and data for similar companies, the Personnel Director of National Industries predicts that the number of employees can be described by the equation

$$N = 200(0.04)^{0.5^t}$$

where N is the number of employees after t years. Assuming he is correct, how many employees will National Industries have after 3 years? How many employees did the company employ initially? How many will it employ when it reaches its maximum size?

The company employs $(200)(0.04) = 8$ people initially and 200 at maximum size. After 3 years it employs

$$N = (200)(0.04)^{0.5^3}$$

$$\log N = \log 200 + 0.5^3 \log 0.04$$

$$= 2.3010 + (0.125)(-1.3979)$$

$$= 2.1263$$

$$N = 133.75 \quad \text{or approximately 134 people}$$

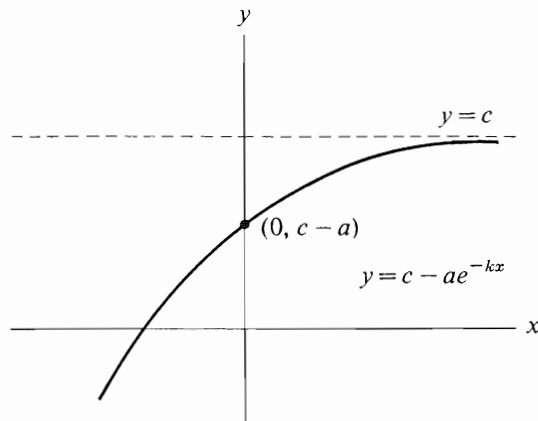
Learning Curves

Because of their extensive use by psychologists to describe learning, exponential curves of the form

$$y = c - ae^{-kx}$$

where c , a , and k are positive, are frequently referred to as learning curves. [Note that this is case (f) of the general form $y = ae^{kx} + c$ discussed previously.] In his reinforcement theory of learning, Clark Hull used the special case $c = a$ of this function as one of the basic equations for describing the relationship between strength of learning y and number of reinforcements x .

Figure 1.100



The function $y = c - ae^{-kx}$ rises rapidly at first, flattens out, and then approaches its asymptote $y = c$ (see Figure 1.100). These curves have been found to be appropriate for representing various cost and production functions.

EXAMPLE

If the number of items y manufactured per day x days after the beginning of a production run is given by

$$y = 200(1 - e^{-0.1x})$$

how many items are manufactured per day 10 days after beginning the run and what per cent of the maximum is this?

$$\begin{aligned} y &= 200(1 - e^{-1}) \\ &= 200(1 - 0.368) \\ &= 200(0.632) \\ &= 126.4 \text{ or approximately 126 items, 63.2\% of the 200 maximum.} \end{aligned}$$

EXAMPLE

If the yearly maintenance cost y of a computer is related to its average monthly use x (in hundreds of hours) by the equation

$$y = 35,000 - 25,000e^{-0.02x}$$

what is the yearly maintenance cost for 200 hours average monthly use?

$$\begin{aligned} y &= 35,000 - 25,000e^{-0.04} \\ &= 35,000 - 25,000(0.135) \\ &= 35,000 - 3375 \\ &= \$31,625 \text{ yearly maintenance} \end{aligned}$$

PROBLEMS

1. Chem Corporation has \$10,000 to deposit and expects to leave the deposit for 20 years. Two options are available: 5% interest payable semiannually and $4\frac{1}{2}\%$ interest payable quarterly. Which option should Chem Corporation choose?
2. In order to have \$20,000 after 20 years at 6% interest payable yearly, how much must be deposited?
3. The National Professional Society has been formed with 10 original members. The rules state that each member may invite 2 people to join at the beginning of each year. If each member always takes advantage of this rule, how many members will the society have after 15 years?
4. Monthly total revenue (in dollars) for a particular firm can be described by the equation

$$R = 1000(0.10)^{0.8p}$$

where p is the amount spent for promotion and advertising. What is total revenue when nothing is spent for promotion and advertising? What is the maximum attainable total revenue? What is total revenue if \$20 is spent for promotion and advertising?

5. Production costs (in hundreds of dollars) for a firm are described by the equation

$$C = 100 - 70e^{-0.02x}$$

where x is the number of units of output. What are the firm's fixed costs? When output is 100 units, what proportion of the production costs are fixed?

6. One job on a production line consists of screwing a small screw into a metal plate. For a typical employee, the number of plates completed per hour is described by the equation

$$y = 50 - 40e^{-0.30x}$$

where x is number of hours the employee has worked on the production line. a. How many plates can an employee complete the first hour? b. How many the sixth hour?

7. The number of firms in a particular industry is described by the equation

$$N = 6(0.5)^{0.75t}$$

where t is the number of years since the industry began. How many firms were there in the industry after 5 years? How many firms were in the industry originally? How many firms will there be when the industry reaches maximum size?

8. In order to have \$10,000 after \$6,000 has been deposited for 25 years with interest payable yearly, what interest rate must be paid?
9. For a 10-year deposit, what interest rate payable annually is equivalent to 5% interest payable quarterly?
10. For a 25-year deposit, what interest payable semiannually is equivalent to 4% payable quarterly?

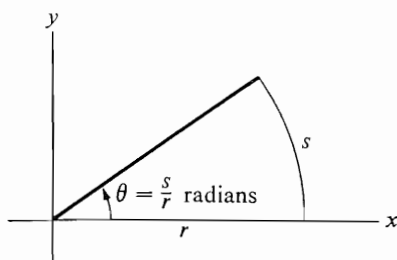
Answers to Odd-Numbered Problems

1. 5% payable semiannually (\$26,841) rather than $4\frac{1}{2}\%$ payable quarterly (\$24,480)
3. 143,489,070
5. \$3000 fixed costs
33.14% of production costs are fixed for 100 units output
7. 3.45 (approximately 3) firms after 5 years
3 firms originally
6 firms at maximum size
9. 5.1%

1.10 Trigonometric Curves

In this section the measurement of angles in degrees and radians is discussed and the six basic trigonometric functions are defined. Some of the basic trigonometric identities are given. The representation of a point in polar coordinates and the transformation of an equation into polar coordinates are also discussed.

Figure 1.101



DEGREE AND RADIAN MEASURE

An angle may be measured in either degrees or radians. Although measurement of angles in degrees is probably more familiar, measurement in radians is more convenient for stating formulas and solving problems in calculus.

An angle is measured in the counterclockwise direction. The measurement of an angle in degrees is based on the arbitrary assumption that a circle contains 360° ; thus a straight line contains 180° and a right angle contains 90° .

The radian or circular measure of an angle whose vertex is at the center of a circle is defined as the ratio of the length of the intercepted arc to the radius of the circle (see Figure 1.101). The unit angle, referred to as a *radian*, is the angle which, when placed with its vertex at the center of a circle, intercepts an arc equal in length to the radius of the circle (see Figure 1.102). Since the circumference of a circle is $2\pi r$, there are 2π radians in a circle, π radians in a straight line, and $\pi/2$ radians in a right angle.

The relationship between radian measure and degree measure may be expressed by

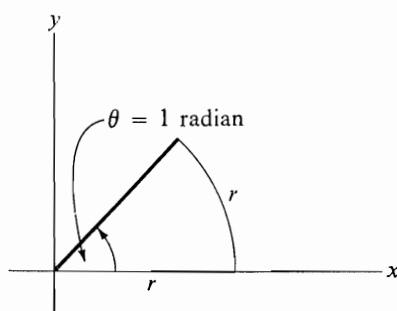
$$\pi \text{ radians} = 180^\circ$$

Thus

$$1 \text{ radian} = \left(\frac{180}{\pi} \right)^\circ \approx 57.296^\circ$$

$$1^\circ = \left(\frac{\pi}{180} \right) \text{ radians} \approx 0.01745 \text{ radians}$$

Figure 1.102



EXAMPLE

$$200^\circ = \frac{200\pi}{180} \text{ radians} = \frac{10\pi}{9} \text{ radians}$$

$$-35^\circ = -\frac{35\pi}{180} \text{ radians} = -\frac{7\pi}{36} \text{ radians}$$

$$300^\circ = \frac{300\pi}{180} \text{ radians} = \frac{5\pi}{3} \text{ radians}$$

TRIGONOMETRIC FUNCTIONS

The trigonometric functions of an angle θ are the sine of θ ($\sin \theta$), the cosine of θ ($\cos \theta$), the tangent of θ ($\tan \theta$), the cosecant of θ ($\csc \theta$), the secant of θ ($\sec \theta$), and the cotangent of θ ($\cot \theta$). When the angle θ is at the center of a circle of radius r and is measured counterclockwise, as in Figure 1.103, the trigonometric functions of θ are defined by the equations

$$\sin \theta = \frac{a}{r} \quad \cos \theta = \frac{b}{r} \quad \tan \theta = \frac{a}{b}$$

$$\csc \theta = \frac{r}{a} \quad \sec \theta = \frac{r}{b} \quad \cot \theta = \frac{b}{a}$$

Note that

$$\csc \theta = \frac{1}{\sin \theta} \quad \sec \theta = \frac{1}{\cos \theta} \quad \cot \theta = \frac{1}{\tan \theta}$$

$$\tan \theta = \frac{\sin \theta}{\cos \theta} \quad \cot \theta = \frac{\cos \theta}{\sin \theta}$$

The signs of the trigonometric functions can be summarized as follows:

$$\begin{array}{c|c} + & + \\ \hline - & - \end{array}$$

sin, csc

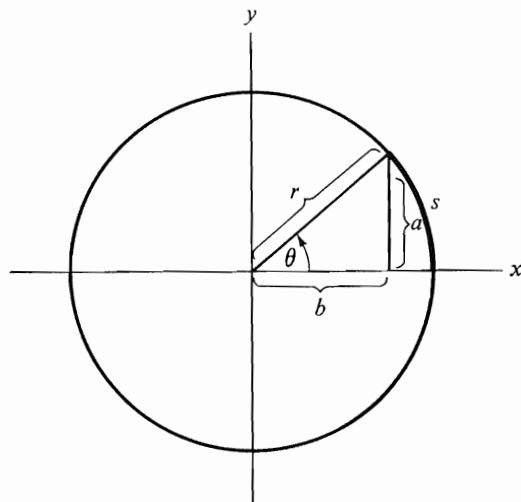
$$\begin{array}{c|c} - & + \\ \hline - & + \end{array}$$

cos, sec

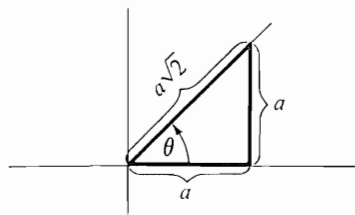
$$\begin{array}{c|c} - & + \\ \hline + & - \end{array}$$

tan, cot

Figure 1.103



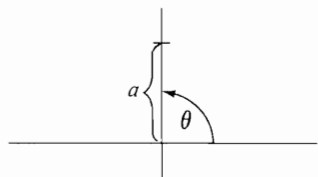
The values of the trigonometric functions can be obtained geometrically for some commonly used angles. For example, if $\theta = 45^\circ = \pi/4$ radians, $b = a$ and, using the Pythagorean theorem, $r^2 = a^2 + b^2$ and $r = a\sqrt{2}$, as shown below.



Thus

$$\sin \theta = \frac{a}{a\sqrt{2}} = \frac{\sqrt{2}}{2} \quad \cos \theta = \frac{a}{a\sqrt{2}} = \frac{\sqrt{2}}{2} \quad \tan \theta = \frac{a}{a} = 1$$

Similarly, if $\theta = 90^\circ = \pi/2$ radians, $r = a$ and $b = 0$, as shown below.



Thus

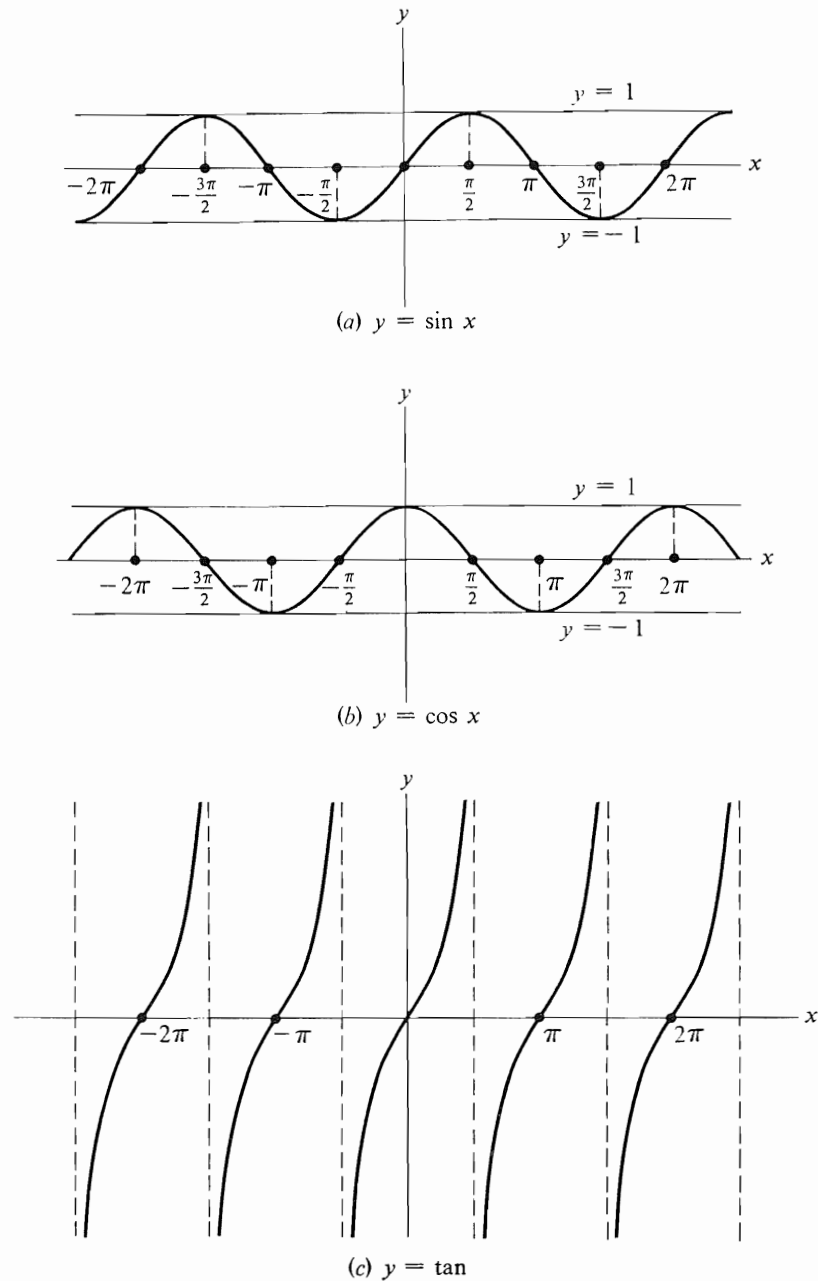
$$\sin \theta = \frac{a}{a} = 1 \quad \cos \theta = \frac{0}{a} = 0 \quad \tan \theta = \frac{a}{0} \text{ not defined}$$

Values of the sine, cosine and tangent are given for some commonly used angles in the table below. These values can all be obtained geometrically.

Degrees	0	30	45	60	90	180	270	360
Radians	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$	π	$\frac{3\pi}{2}$	2π
Sine	0	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1	0	-1	0
Cosine	1	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{1}{2}$	0	-1	0	1
Tangent	0	$\frac{\sqrt{3}}{3}$	1	$\sqrt{3}$	not defined	0	not defined	0

The graphs of the sine, cosine, and tangent functions are given in Figure 1.104. Note that the function $\sin x$ is defined and continuous for all values of x . $\sin x$ is a periodic function with the period 2π , since $\sin(x + 2\pi) = \sin x$ for any x , that is, when the value of x is increased by 2π , the value of y is unchanged. The function $\cos x$

Figure 1.104



is also defined and continuous for all values of x and is periodic with period 2π . The graph of $y = \cos x$ can be obtained from the graph of $y = \sin x$ by taking the line $x = \pi/2$ as the y -axis; that is, the graph of $y = \cos x$ is the same as the graph of $y = \sin x$ when that graph is shifted to the right by $\pi/2$ radians. The function $\tan x$ is discontinuous for all values of x such that $x = (n + \frac{1}{2})\pi$, where n is any positive or negative integer; $\tan x$ is periodic with period π .

The symmetry and periodicity of the sine, cosine, and tangent functions make it possible to obtain a trigonometric function of any angle, $0 \leq \theta \leq 2\pi$, from the trigonometric function of an angle θ such that $0 \leq \theta \leq \pi/2$. Thus trigonometric functions are tabled only for angles between 0 and $\pi/2$. The formulas for reducing angles are given below. Note, for example, that $\sin(-\theta) = -\sin \theta$, $\cos(-\theta) = \cos \theta$, $\tan(-\theta) = -\tan \theta$, and so forth.

Angle	sin	cos	tan	cot	sec	csc
$-\theta$	$-\sin \theta$	$\cos \theta$	$-\tan \theta$	$-\cot \theta$	$\sec \theta$	$-\csc \theta$
$\frac{\pi}{2} - \theta$	$\cos \theta$	$\sin \theta$	$\cot \theta$	$\tan \theta$	$\csc \theta$	$\sec \theta$
$\frac{\pi}{2} + \theta$	$\cos \theta$	$-\sin \theta$	$-\cot \theta$	$-\tan \theta$	$-\csc \theta$	$\sec \theta$
$\pi - \theta$	$\sin \theta$	$-\cos \theta$	$-\tan \theta$	$-\cot \theta$	$-\sec \theta$	$\csc \theta$
$\pi + \theta$	$-\sin \theta$	$-\cos \theta$	$\tan \theta$	$\cot \theta$	$-\sec \theta$	$-\csc \theta$
$\frac{3\pi}{2} - \theta$	$-\cos \theta$	$-\sin \theta$	$\cot \theta$	$\tan \theta$	$-\csc \theta$	$-\sec \theta$
$\frac{3\pi}{2} + \theta$	$-\cos \theta$	$\sin \theta$	$-\cot \theta$	$-\tan \theta$	$\csc \theta$	$-\sec \theta$
$2\pi - \theta$	$-\sin \theta$	$\cos \theta$	$-\tan \theta$	$-\cot \theta$	$\sec \theta$	$-\csc \theta$

EXAMPLE

Find the value of each of the following trigonometric functions: $\cos 3\pi/4$, $\tan 5\pi/6$, $\sin 5\pi/3$, $\sec 7\pi/6$, $\csc 2\pi/3$, $\cot(-\pi/4)$.

$$\cos \frac{3\pi}{4} = \cos \left(\frac{\pi}{2} + \frac{\pi}{4} \right) = -\sin \frac{\pi}{4} = -\frac{\sqrt{2}}{2}$$

$$\tan \frac{5\pi}{6} = \tan \left(\pi - \frac{\pi}{6} \right) = -\tan \frac{\pi}{6} = -\frac{\sqrt{3}}{3}$$

$$\sin \frac{5\pi}{3} = \sin \left(2\pi - \frac{\pi}{3} \right) = -\sin \frac{\pi}{3} = -\frac{\sqrt{3}}{2}$$

$$\sec \frac{7\pi}{6} = \sec \left(\pi + \frac{\pi}{6} \right) = -\sec \frac{\pi}{6} = -\frac{1}{\cos \frac{\pi}{6}} = -\frac{2\sqrt{3}}{3}$$

$$\csc \frac{2\pi}{3} = \csc \left(\pi - \frac{\pi}{3} \right) = \csc \frac{\pi}{3} = \frac{1}{\sin \frac{\pi}{3}} = \frac{2\sqrt{3}}{3}$$

$$\cot \left(-\frac{\pi}{4} \right) = -\cot \frac{\pi}{4} = -\frac{1}{\tan \frac{\pi}{4}} = -1$$

TRIGONOMETRIC IDENTITIES

There are a number of trigonometric identities which are useful in simplifying results involving trigonometric functions. Some of the more commonly used identities are given below without proof; additional identities and proofs can be found in any trigonometry book.

$$\sin^2 \theta + \cos^2 \theta = 1 \quad \sec^2 \theta - \tan^2 \theta = 1 \quad \csc^2 \theta - \cot^2 \theta = 1$$

$$\sin (x + y) = \sin x \cos y + \cos x \sin y$$

$$\sin (x - y) = \sin x \cos y - \cos x \sin y$$

$$\cos (x + y) = \cos x \cos y - \sin x \sin y$$

$$\cos (x - y) = \cos x \cos y + \sin x \sin y$$

$$\tan (x + y) = \frac{\tan x + \tan y}{1 - \tan x \tan y} \quad \tan (x - y) = \frac{\tan x - \tan y}{1 + \tan x \tan y}$$

$$\sin 2x = 2 \sin x \cos x \quad \cos 2x = \cos^2 x - \sin^2 x \quad \tan 2x = \frac{2 \tan x}{1 - \tan^2 x}$$

$$\sin \frac{x}{2} = \pm \sqrt{\frac{1 - \cos x}{2}} \quad \cos \frac{x}{2} = \pm \sqrt{\frac{1 + \cos x}{2}}$$

$$\tan \frac{x}{2} = \pm \sqrt{\frac{1 - \cos x}{1 + \cos x}}$$

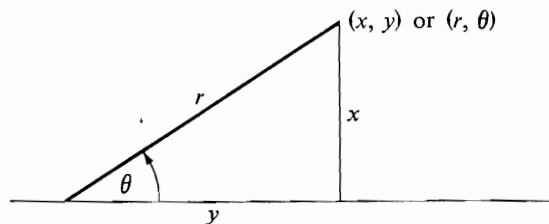
$$\sin^2 x = \frac{1}{2} - \frac{1}{2} \cos 2x \quad \cos^2 x = \frac{1}{2} + \frac{1}{2} \cos 2x$$

POLAR COORDINATES

It is frequently more convenient to describe a point by its polar coordinates rather than by its rectangular coordinates. If a point is described by the rectangular coordinates (x, y) , then it can also be described by the *polar coordinates* (r, θ) as follows. Consider a line drawn between the point and a given origin; the angle θ which this line makes with the horizontal axis and the distance r between the point and the origin determine the point, as shown in Figure 1.105.

Note that the polar coordinates of a point are not unique. For example, the polar coordinates (r, θ) , $(r, \theta + 2\pi)$, and $(r, \theta - 2\pi)$ all represent the same point. In practice, the polar coordinates of a point are usually written using the angle between 0 and 2π .

Figure 1.105



The rectangular coordinates (x, y) and the polar coordinates (r, θ) of a point are related as follows:

$$x = r \cos \theta \qquad y = r \sin \theta \qquad r = \sqrt{x^2 + y^2}$$

$$\sin \theta = \frac{y}{\sqrt{x^2 + y^2}} \qquad \cos \theta = \frac{x}{\sqrt{x^2 + y^2}} \qquad \tan \theta = \frac{y}{x}$$

These relations are used to change from one system of coordinates to the other.

EXAMPLE

Find polar coordinates for the following points given in rectangular coordinates,

(a) $(0, \sqrt{2})$

(b) $(1, 1)$

(a) $(x, y) = (0, \sqrt{2})$

$$r = \sqrt{x^2 + y^2} = \sqrt{0 + 2} = \sqrt{2}$$

$$\cos \theta = 0 \quad \text{and} \quad \sin \theta = 1$$

so

$$\theta = \frac{\pi}{2}$$

Polar coordinates $(r, \theta) = \left(\sqrt{2}, \frac{\pi}{2} \right)$.

(b) $(x, y) = (1, 1)$

$$r = \sqrt{x^2 + y^2} = \sqrt{1 + 1} = \sqrt{2}$$

$$\cos \theta = \frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2} \quad \text{and} \quad \sin \theta = \frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2}$$

so

$$\theta = \frac{\pi}{4}$$

Polar coordinates $(r, \theta) = \left(\sqrt{2}, \frac{\pi}{4} \right)$.

EXAMPLE

Find rectangular coordinates for the following points given in polar coordinates.

(a) $\left(3, \frac{\pi}{2} \right)$

(b) $\left(\sqrt{2}, \frac{\pi}{6} \right)$

$$(a) \quad (r, \theta) = \left(3, \frac{\pi}{2}\right)$$

$$r \cos \theta = (3)(0) = 0$$

$$r \sin \theta = (3)(1) = 3$$

$$(x, y) = (0, 3)$$

$$(b) \quad (r, \theta) = \left(\sqrt{2}, \frac{\pi}{6}\right)$$

$$r \cos \theta = (\sqrt{2})\left(\frac{\sqrt{3}}{2}\right) = \frac{\sqrt{6}}{2}$$

$$r \sin \theta = (\sqrt{2})\left(\frac{1}{2}\right) = \frac{\sqrt{2}}{2}$$

$$(x, y) = \left(\frac{\sqrt{6}}{2}, \frac{\sqrt{2}}{2}\right)$$

EXAMPLE

Find the equation in rectangular coordinates of the curve whose equation in polar coordinates is $r = 4/(1 + 2 \cos \theta)$.

$$\sqrt{x^2 + y^2} = \frac{4}{1 + \frac{2x}{\sqrt{x^2 + y^2}}}$$

$$\sqrt{x^2 + y^2} + 2x = 4$$

$$\sqrt{x^2 + y^2} = 4 - 2x$$

$$x^2 + y^2 = 16 - 16x + 4x^2$$

$$3x^2 - y^2 - 16x + 16 = 0 \quad (\text{hyperbola})$$

EXAMPLE

Find the equation in polar coordinates of the curve whose equation in rectangular coordinates is $Ax + By + C = 0$.

$$Ar \cos \theta + Br \sin \theta + C = 0$$

$$r(A \cos \theta + B \sin \theta) + C = 0$$

PROBLEMS

1. Find the number of radians in each of the following angles given in degrees.

a. $\theta = -240^\circ$

d. $\theta = -40^\circ$

b. $\theta = 70^\circ$

e. $\theta = -100^\circ$

c. $\theta = 120^\circ$

f. $\theta = 315^\circ$

2. Find the value of the following trigonometric functions.

a. $\sin \frac{7\pi}{6}$

b. $\csc \frac{4\pi}{3}$

c. $\tan \frac{5\pi}{4}$

d. $\cos \frac{5\pi}{3}$

3. Obtain polar coordinates for the following points given in rectangular coordinates.

a. $(1, \sqrt{3})$

c. $(3, 3\sqrt{3})$

b. $(-1, -1)$

d. $(-1, 0)$

4. Obtain rectangular coordinates for the following points given in polar coordinates.

a. $(4, \frac{\pi}{3})$

c. $(3, \frac{\pi}{6})$

b. $(-2, \pi)$

d. $(-1, \frac{\pi}{2})$

Obtain the equation in rectangular coordinates that corresponds to each of the following equations given in polar coordinates.

5. $r = 5$

12. $r = a \sec \theta$

6. $r = \sin \theta + \cos \theta$

13. $r^2 = a^2 \cot \theta$

7. $r = \frac{2}{2 - \cos \theta}$

14. $r = \tan \theta \sec \theta$

8. $r = 1 + \cos \theta$

15. $r^3 = a^3 \csc \theta$

9. $r = a \cos \theta$

16. $r = a \tan^2 \theta$

10. $r = a \sin \theta$

17. $r^2 = \theta$

11. $r = \theta$

18. $\theta = \frac{\pi}{2}$

Obtain the equation in polar coordinates that corresponds to each of the following equations given in rectangular coordinates.

19. $y^2 = 4ax$

24. $(x^2 + y^2)^2 = a^2(x^2 - y^2)$

20. $x^2 + y^2 = a^2$

25. $2xy = a^2$

21. $x^2 + y^2 = x + y$

26. $y = x^2$

22. $y = x$

27. $xy^2 = a$

23. $y^3 = ax^2$

28. $x + y = a$

Answers to Odd-Numbered Problems

1. a. $s = -\frac{4\pi}{3}$

b. $(\sqrt{2}, \frac{5\pi}{4})$

b. $s = \frac{7\pi}{18}$

c. $(6, \frac{\pi}{3})$

c. $s = \frac{2\pi}{3}$

d. $(1, \pi)$

d. $s = -\frac{2\pi}{9}$

5. $x^2 + y^2 = 25$

7. $3x^2 + 4y^2 - 4x - 4 = 0$

e. $s = -\frac{5\pi}{9}$

9. $x^2 + y^2 - ax = 0$

11. $y = x \tan(\sqrt{x^2 + y^2})$

f. $s = \frac{7\pi}{4}$

13. $y^3 + x^2y - a^2x = 0$

15. $x^2y + y^3 = a^3$

17. $y = x \tan(x^2 + y^2)$

19. $r = 4a \cot \theta \csc \theta$

21. $r = \cos \theta + \sin \theta$

3. a. $(2, \frac{\pi}{3})$

23. $r = a \cot^2 \theta \csc \theta$

25. $r^2 = \frac{1}{2}a^2 \sec \theta \csc \theta$

27. $r^3 = a \sec \theta \csc^2 \theta$

1.11 Applications of Trigonometric Curves in Business and Economics

Many economic phenomena can be described in part by sine or cosine functions. In the study of business cycles, seasonal or other cyclic variations are described by sine or cosine functions. In spectral analysis, the method most frequently used currently to analyze periodicity or cyclic variation in time series, the detrended series is analyzed in terms of cosine functions. Spectral analysis, and even less complicated types of analysis of cyclic variation, involve statistical problems beyond the scope of the present discussion. However, these analyses are an application of trigonometric curves, and the logic on which they are based is intuitively reasonable: A series is described in terms of components representing trends and cyclic variations.

Technical Note I Conditions for Equilibrium

Two linear equations of the form

$$A_1x + B_1y + C_1 = 0 \quad A_2x + B_2y + C_2 = 0$$

have the simultaneous solution (x, y) , where

$$x = \frac{B_1C_2 - B_2C_1}{A_1B_2 - A_2B_1} \quad y = \frac{A_2C_1 - A_1C_2}{A_1B_2 - A_2B_1}$$

provided $A_1B_2 - A_2B_1 \neq 0$.

Thus if the nonvertical demand and supply curves

$$y = m_Dx + b_D \quad y = m_Sx + b_S$$

(where the subscripts D and S distinguish demand and supply curves) are written

$$m_Dx - y + b_D = 0 \quad m_Sx - y + b_S = 0$$

then the equilibrium quantity and price are

$$x = \frac{b_D - b_S}{m_S - m_D} \quad y = \frac{b_Dm_S - b_Sm_D}{m_S - m_D}$$

since $m_S - m_D \neq 0$. Since $m_D < 0$, $m_S > 0$, $y_D > 0$, and y_S is unrestricted, $m_S - m_D > 0$ and $x \geq 0$ if $b_D - b_S \geq 0$, that is, if $b_D \geq b_S$, which is the first condition stated for equilibrium in the first quadrant; $y \geq 0$ if

$$b_Dm_S - b_Sm_D \geq 0$$

or

$$b_Dm_S \geq b_Sm_D$$

$$\frac{b_D}{m_D} \leq \frac{b_S}{m_S} \quad (\text{since } m_D < 0)$$

$$\frac{-b_D}{m_D} \geq \frac{-b_S}{m_S}$$

which is the second condition stated for equilibrium in the first quadrant, since the x -coordinate of the x -intercept of the line $y = mx + b$ is $-(b/m)$.

Technical Note II Identification of Degenerate and Imaginary Loci

The type of conic section represented by a second-degree equation, $Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$, can be determined from certain conditions on the coefficients as summarized in Table 1.1.

Table 1.1

Case	Conditions on the coefficients	Type of conic
Proper conic, $\bar{D} \neq 0$	$\bar{C} > 0$; I, \bar{D} opposite in sign, $A = C$, $B = 0$ (no xy -term)	Circle
	$\bar{C} > 0$; I, \bar{D} opposite in sign	Ellipse
	$\bar{C} < 0$	Hyperbola
	$\bar{C} = 0$	Parabola
	$\bar{C} > 0$, I, \bar{D} same in sign	No real locus
Degenerate conic, $\bar{D} = 0$	$\bar{C} < 0$	Two intersecting lines
	$\bar{C} = 0, J < 0$	Two parallel lines
	$\bar{C} = 0, J = 0$	Two coincident lines
	$\bar{C} > 0$	A point
	$\bar{C} = 0, J > 0$	No real locus

In this table,

$$I = A + C$$

$$\bar{C} = AC - \frac{B^2}{4}$$

$$J = AC + CF + AF - \frac{E^2}{4} - \frac{D^2}{4} - \frac{B^2}{4}$$

$$\bar{D} = ACF + \frac{BED}{4} - \frac{CD^2}{4} - \frac{AE^2}{4} - \frac{FB^2}{4}$$

In many cases it can be assumed (owing to the nature of the application) that the given equation represents a proper conic (that is, $\bar{D} \neq 0$) and a real locus (that is, I and \bar{D} opposite in sign if $\bar{C} > 0$). In such cases it is necessary only to compute \bar{C} and note the coefficients A , C , and B in order to determine which type of conic section is represented by the equation.

Technical Note III Completing the Square

In many problems it is convenient to have an equation written so that variables which occur in terms of second degree occur only as terms of a perfect square; this is accomplished by putting the terms of the perfect square(s) on one side and all other terms on the other side of the equation and adding the appropriate constant terms to both sides of the equation.

The process of completing the square is based on the formula $(x + a)^2 = x^2 + 2ax + a^2$. Thus, given $x^2 + cx$, take half the coefficient of x and square it to complete the square: $x^2 + cx + (c/2)^2$. Any constant term added to one side of an equation must, of course, be added also to the other side in order to preserve the equality. Completing the square is the essential step in putting a second-degree equation into the appropriate standard form.

If $ax^2 + bx + cy + d = 0$, then the terms to be included in the perfect square are ax^2 and bx ; thus

$$\begin{aligned} a\left(x^2 + \frac{b}{a}x\right) &= -cy - d \\ a\left(x^2 + \frac{b}{a}x + \frac{b^2}{4a^2}\right) &= -cy - d + \frac{b^2}{4a} \\ \left(x + \frac{b}{2a}\right)^2 &= -\frac{c}{a}\left(y + \frac{d}{c} - \frac{b^2}{4ac}\right) \end{aligned}$$

EXAMPLE

Complete the square for the equation

$$\begin{aligned} 3x^2 + 4x + 2y - 4 &= 0 \\ 3\left(x^2 + \frac{4}{3}x + \frac{4}{9}\right) &= -2y + 4 + \frac{4}{3} \\ \left(x + \frac{2}{3}\right)^2 &= -\frac{2}{3}\left(y - \frac{4}{3}\right) \end{aligned}$$

If $ax^2 + by^2 + cx + dy + e = 0$, then the terms to be included in the perfect squares are ax^2 , cx , by^2 , and dy ; thus

$$\begin{aligned} a\left(x^2 + \frac{c}{a}x\right) + b\left(y^2 + \frac{d}{b}y\right) &= -e \\ a\left(x^2 + \frac{c}{a}x + \frac{c^2}{4a^2}\right) + b\left(y^2 + \frac{d}{b}y + \frac{d^2}{4b^2}\right) &= -e + \frac{c^2}{4a} + \frac{d^2}{4b} \\ \frac{\left(x + \frac{c}{2a}\right)^2}{b} + \frac{\left(y + \frac{d}{2b}\right)^2}{a} &= \frac{1}{ab}\left(\frac{c^2}{4a} + \frac{d^2}{4b} - e\right) = \frac{c^2}{4a^2b} + \frac{d^2}{4ab^2} - \frac{e}{ab} \end{aligned}$$

EXAMPLE

Complete the squares for the equation

$$\begin{aligned} 3x^2 + 4y^2 - 6x + 7y - 10 &= 0 \\ 3(x^2 - 2x) + 4(y^2 + \frac{7}{4}y) &= 10 \\ 3(x^2 - 2x + 1) + 4(y^2 + \frac{7}{4}y + \frac{49}{64}) &= 10 + 3 + \frac{49}{16} \\ \frac{(x-1)^2}{4} + \frac{(y+\frac{7}{8})^2}{3} &= \frac{1}{12}(13 + \frac{49}{16}) = \frac{257}{192} \end{aligned}$$

The essential aspect of the process of completing the square is the addition of the appropriate constant terms so that groups of terms form perfect squares; the most convenient form of writing the constant coefficients depends on the particular problem.

A second-degree equation can be put into the appropriate standard form for the type of conic it represents by the process of completing the square.

The circle:

$$Ax^2 + Ay^2 + Dx + Ey + F = 0$$

$$A\left(x^2 + \frac{D}{A}x + \frac{D^2}{4A^2}\right) + A\left(y^2 + \frac{E}{A}y + \frac{E^2}{4A^2}\right) = \frac{D^2}{4A} + \frac{E^2}{4A} - F$$

$$\left(x + \frac{D}{2A}\right)^2 + \left(y + \frac{E}{2A}\right)^2 = \frac{D^2}{4A^2} + \frac{E^2}{4A^2} - \frac{F}{A}$$

which is of the form $(x - h)^2 + (y - k)^2 = r^2$, where

$$h = -\frac{D}{2A} \quad k = -\frac{E}{2A} \quad r^2 = \frac{D^2}{4A^2} + \frac{E^2}{4A^2} - \frac{F}{A}$$

The ellipse: $Ax^2 + Cy^2 + Dx + Ey + F = 0$, A and C same sign

$$A\left(x^2 + \frac{D}{A}x + \frac{D^2}{4A^2}\right) + C\left(y^2 + \frac{E}{C}y + \frac{E^2}{4C^2}\right) = \frac{D^2}{4A} + \frac{E^2}{4C} - F$$

Denoting $\frac{D^2}{4A} + \frac{E^2}{4C} - F$ by k ,

$$\frac{\left(x + \frac{D}{2A}\right)^2}{\frac{k}{A}} + \frac{\left(y + \frac{E}{2C}\right)^2}{\frac{k}{C}} = 1$$

which is of the form

$$\frac{(x - h)^2}{a^2} + \frac{(y - k)^2}{b^2} = 1$$

where

$$h = -\frac{D}{2A} \quad k = -\frac{E}{2C}$$

The parabola: $Ax^2 + Dx + Ey + F = 0$

$$A\left(x^2 + \frac{D}{A}x + \frac{D^2}{4A^2}\right) = -E\left(y + \frac{F}{E} - \frac{D^2}{4AE}\right)$$

$$\left(x + \frac{D}{2A}\right)^2 = -\frac{E}{A}\left(y - \frac{D^2 - 4AF}{4AE}\right)$$

which is of the form $(x - h)^2 = 4p(y - k)$, where

$$h = -\frac{D}{2A} \quad k = \frac{D^2 - 4AF}{4AE} \quad 4p = -\frac{E}{A}$$

(and similarly for the parabola $Cy^2 + Dx + Ey + F = 0$).

The hyperbola: $Ax^2 + Cy^2 + Dx + Ey + F = 0$, A and C opposite in sign

$$A\left(x^2 + \frac{D}{A}x + \frac{D^2}{4A^2}\right) + C\left(y^2 + \frac{E}{C}y + \frac{E^2}{4C^2}\right) = \frac{D^2}{4A} + \frac{E^2}{4C} - F$$

Denoting $\frac{D^2}{4A} + \frac{E^2}{4C} - F$ by k ,

$$\frac{\left(x + \frac{D}{2A}\right)^2}{\frac{k}{A}} + \frac{\left(y + \frac{E}{2C}\right)^2}{\frac{k}{C}} = 1$$

which is of the form

$$\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = 1$$

or

$$\frac{(y-k)^2}{b^2} - \frac{(x-h)^2}{a^2} = 1$$

since A and C are opposite in sign, where

$$h = -\frac{D}{2A} \quad k = -\frac{E}{2C}$$

The various properties of particular second-degree curves can, of course, be obtained by identifying the appropriate type of conic section and then substituting appropriate numerical values in the corresponding formulas given above, rather than by putting each equation into standard form.

Technical Note IV Formula for Roots of a Quadratic Equation

The formula for the roots of a quadratic equation is obtained by completing the square as follows:

$$ax^2 + bx + c = 0$$

$$a\left(x^2 + \frac{bx}{a} + \frac{b^2}{4a^2}\right) = -c + \frac{b^2}{4a}$$

$$\left(x + \frac{b}{2a}\right)^2 = \frac{b^2 - 4ac}{4a^2}$$

$$x + \frac{b}{2a} = \pm \sqrt{\frac{b^2 - 4ac}{4a^2}}$$

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Technical Note V x-intercepts of $y = ae^{kx} + c$ and $y = A \ln(1+x) + B$

To obtain the x -intercept for the curve $y = ae^{kx} + c$: Let $y = 0$; then

$$ae^{kx} = -c$$

$$e^{kx} = -\frac{c}{a}$$

$$\ln\left(-\frac{c}{a}\right) = kx$$

$$x = \frac{1}{k} \ln\left(-\frac{c}{a}\right)$$

In order for $\ln\left(-\frac{c}{a}\right)$ to be defined,

$$c \leq 0, a > 0 \quad [\text{cases (c) and (d)}]$$

or

$$c > 0, a < 0 \quad [\text{cases (e) and (f)}]$$

Note in Figure 1.94 that only cases (c), (d), (e), and (f) have x -intercepts and that

$$\frac{1}{k} \ln \left(-\frac{c}{a} \right) > 0 \text{ if } a > 0, c \leq 0, k < 0, a > |c| \quad [\text{case (d)}]$$

or

$$a < 0, c > 0, k > 0, c > |a| \quad [\text{case (e)}]$$

$$\frac{1}{k} \ln \left(-\frac{c}{a} \right) < 0 \text{ if } a > 0, c \leq 0, k > 0, a > |c| \quad [\text{case (c)}]$$

or

$$a < 0, c > 0, k < 0, c > |a| \quad [\text{case (f)}]$$

To obtain the x -intercept for the curve $y = A \ln(1+x) + B$: Let $y = 0$; then

$$A \ln(1+x) = -B$$

$$\ln(1+x) = -\frac{B}{A}$$

$$e^{-(B/A)} = 1+x$$

$$x = e^{-(B/A)} - 1$$

and

$$e^{-(B/A)} - 1 > 0 \quad \text{if } \frac{B}{A} < 0$$

$$e^{-(B/A)} - 1 = 0 \quad \text{if } \frac{B}{A} = 0 \quad \text{that is, if } B = 0$$

$$e^{-(B/A)} - 1 < 0 \quad \text{if } \frac{B}{A} > 0$$

Thus

$$e^{-(B/A)} - 1 > 0$$

if

$$A > 0, B > 0 \quad [\text{case (a)}]$$

$$A < 0, B < 0 \quad [\text{case (d)}]$$

and

$$e^{-(B/A)} - 1 < 0$$

if

$$A > 0, B < 0 \quad [\text{case (c)}]$$

$$A < 0, B > 0 \quad [\text{case (b)}]$$

Differential Calculus: Functions of One Variable

2

2.1 Introduction

Calculus is concerned with the mathematical analysis of change or movement. Because everything in the world changes, calculus has applications in virtually all areas of scientific inquiry. It is nearly impossible to exaggerate the importance of calculus, particularly differential calculus, as a basis for mathematical analysis.

As a distinctly new mathematical method, calculus was developed in the seventeenth century by Sir Isaac Newton and Gottfried Leibnitz, working independently. For Newton, calculus originated in attempts to solve certain problems connected with his work in physics and astronomy: finding the velocity of a moving body, the work done by force, the center of mass of a body. For Leibnitz, calculus originated in attempts to solve certain problems in geometry: finding the tangent to a curve, the length of a portion of a curve, the area bounded by one or more curves, the volume of a solid.

The basic operations of calculus are differentiation and integration; these operations are the inverses of each other, as are addition and subtraction, and multiplication and division. Differentiation is concerned essentially with determining the rate of change of a given function. Integration is concerned essentially with the inverse problem of finding a function when its rate of change is given.

The analogy of a moving-picture film is frequently used in discussing the processes of differentiation and integration. A moving-picture film is a series of (static) pictures, each at least slightly different from the others; each frame depicts the subjects in given positions at a particular instant of time. When the film is run through a projector at proper speed, the pictures are summed up and the illusion of motion is created. Similarly, differentiation essentially breaks up a function into many infinitesimally small (static) pieces and thus analyzes it at a particular point of time or for a particular value of the independent variable; integration, on the other hand, sums up the infinitesimally small pieces to obtain the function.

When relations among variables are stated in equations, calculus can be used to analyze these relations. Calculus has been used by physicists, astronomers, chemists, and engineers almost since its development; more recently, calculus has also been used by biologists, sociologists, psychologists, and economists.

Since analysis in business and economics is frequently concerned with change,

calculus is an extremely valuable tool for these problems. Marginal analysis is perhaps the most direct application of calculus in business and economics; marginal rate of change or variation on the margin is expressed analytically as the first derivative of the relevant function. Differential calculus is also the method by which maxima or minima of functions are obtained. Thus problems of maximizing profit or minimizing cost under various assumptions can be solved using calculus. Mathematical programming, which is concerned with maximizing or minimizing a function subject to constraints, is applied increasingly in business and economics; the methods used in mathematical programming are based on differential calculus.

The idea of the rate of change of a function, which is the basis of differential calculus, is discussed in the following paragraphs. The applications of this concept to marginal analysis and to various problems of maximization and minimization are discussed in some detail in following chapters.

The simplest type of functional relationship between two variables is represented by a straight line and corresponds to a *constant*, or *uniform*, rate of change in the dependent variable with respect to change in the independent variable. A *variable* rate of change in the dependent variable with respect to change in the independent variable is represented by a curvilinear (or nonlinear) function. The *average variable* rate of change is the average value over an interval of the variable rate of change.

For many analyses the most important concept is that of *instantaneous* rate of change, that is, the variable rate of change at a particular instant of time or for a particular value of the independent variable. Instantaneous rate of change is obtained by differentiation and is the first derivative of the function evaluated at the point of interest. The concept of instantaneous change is the basis of marginal analysis in economics; marginal analysis considers the effect on the dependent variable of small changes in the independent variable, that is, variation on the margin.

The mathematical definition and derivation of instantaneous or marginal rate of change are discussed in detail later; the concept can perhaps best be understood intuitively in terms of an example of physical movement. If a car is driven from city *A* to city *B* at a constant speed, then the rate of change in its distance from city *A* is constant with respect to change in time since leaving city *A*; the car is traveling at a constant or uniform rate. However, if the car occasionally slows down for curves, towns, traffic, and so forth, it travels at a variable rate. Suppose the trip from city *A* to city *B* takes 5 hours; the number of miles traveled each hour could be averaged to obtain the average variable rate of travel per hour. There is clearly another rate of travel of interest to both drivers and traffic policemen—the rate of travel at a particular instant of time. It is this instantaneous rate that is obtained by differential calculus.

Although rate of change is perhaps most readily understood in terms of physical motion, it can be generalized to any type of functional relationship. For example, total cost is a function of the quantity produced and usually changes at a variable rate as the quantity produced changes. The rate of change of total cost as quantity produced changes is the marginal cost and is the first derivative of total cost; marginal cost is a function of the quantity produced and can be evaluated for any quantity of interest.

The first derivative, that is, the rate of change of a function, can be used to determine its maximum and minimum points, if any. A function increases (positive rate of change) until it reaches a maximum and then decreases (negative rate of change); similarly, a function decreases until it reaches a minimum and then increases. This basic method for determining the maxima and minima of a function by differentiation has been developed and generalized for use in problems of varying complexity; as noted above, it is the basis of the methods of mathematical programming.

Calculus is concerned with infinitesimally small changes in the independent and dependent variables. Mathematically, such changes are defined using the concepts of limits and continuity. Thus the first sections of this chapter concern the mathematical concepts of limits and continuity, which provide the foundation for the theory of calculus.

The first derivative of a function is then defined and some of its applications in business and economics are discussed. Rules for differentiation of various types of functions are given. Differentials and their use in approximating functions, higher-order derivatives, implicit differentiation, and differentiability and continuity are also discussed. The applications of first and second derivatives in graphical representation are discussed and illustrated. A number of applications of derivatives in business and economics are considered. Indeterminate forms are defined and discussed. Convergence of sequences and series is considered and the use of Taylor's theorem for approximating functions is illustrated.

2.2 Limits

The concept of a limit seems to be one of the most troublesome in mathematics. Apparently, the idea of approaching a point or value arbitrarily closely and still never reaching it is not intuitively appealing. Actually, limit-type concepts are frequently used in nonmathematical thinking and conversation. For example, the theoretical maximum production of a machine or of a factory is a limit, the ideal (or limiting) performance which is never attained in practice but which can be approached arbitrarily closely. This same idea applies to the performance of any mechanical or electronic device for which engineers can calculate an ideal (or limiting) performance; it applies also, for example, to profits under ideal conditions, gas mileage under ideal conditions and operation, and so forth. Similarly, there are lower limits of cost, waste, spoilage, and so forth.

The mathematical concept of a limit is fundamental in understanding differential calculus and is discussed in some detail below. Properties of limits are also considered; these properties are used in later sections concerning series and indeterminate forms.

DEFINITION OF A LIMIT

Consider a function $f(x)$ and let the independent variable x assume values near a given constant a ; then the function $f(x)$ assumes a corresponding set of values. Suppose that when x is close to a , the corresponding values of $f(x)$ are close to some constant A . Moreover, suppose that the values of $f(x)$ can be made to differ arbitrarily little from A by taking values of x that are sufficiently close to a , but not equal to a and that this is true for all such values of x . Then $f(x)$ is said to approach the *limit* A as x approaches a . More concisely, the definitions of the limit of a variable and the limit of a function are as follows.

A variable x is said to approach a constant a as a limit when x varies in such a way that the absolute difference $|x - a|$ becomes and remains less than any preassigned positive number, however small this number is chosen. This is indicated by the notation

$$\lim x = a \quad \text{or} \quad x \rightarrow a$$

EXAMPLE

If x assumes the sequence of values

$$\frac{1}{2}, \frac{3}{4}, \frac{7}{8}, \frac{15}{16}, \dots, \frac{2^n - 1}{2^n}, \dots$$

then $x \rightarrow 1$. But if x assumes the sequence of values

$$\frac{1}{2}, -\frac{3}{4}, \frac{7}{8}, -\frac{15}{16}, \dots, (-1)^{n-1} \frac{2^n - 1}{2^n}, \dots$$

then x does not approach a limit.

If the function $f(x)$ approaches a constant A when x approaches a in whatever manner without assuming the value a , A is said to be the limit of $f(x)$ as x approaches a . This is indicated by the notation

$$\lim_{x \rightarrow a} f(x) = A \quad \text{or} \quad f(x) \rightarrow A \text{ as } x \rightarrow a$$

EXAMPLE

If $f(x) = 2x + 5$, $\lim_{x \rightarrow 0} f(x) = 5$, since, for example,

$$f(1) = 7 \qquad f(-1) = 3$$

$$f\left(\frac{1}{2}\right) = 6 \qquad f\left(-\frac{1}{2}\right) = 4$$

$$f\left(\frac{1}{4}\right) = 5\frac{1}{2} \qquad f\left(-\frac{1}{4}\right) = 4\frac{1}{2}$$

$$f\left(\frac{1}{100}\right) = 5\frac{1}{50} \qquad f\left(-\frac{1}{100}\right) = 4\frac{49}{50}$$

$$f\left(\frac{1}{1000}\right) = 5\frac{1}{500} \qquad f\left(-\frac{1}{1000}\right) = 4\frac{499}{500}$$

etc.

etc.

The following two statements of the definition of the limit of a function are equivalent to the definition given above.

A function $f(x)$ is said to approach a limit A as x approaches a if the absolute difference between $f(x)$ and A is less than an arbitrarily small positive number for all values of x that are sufficiently close to a and for which $x \neq a$.

A function $f(x)$ approaches a limit A as x approaches a if, and only if, for each $\varepsilon > 0$ there exists a δ such that whenever $0 < |x - a| < \delta$, $|f(x) - A| < \varepsilon$.

It is implied in the above discussion of limits that both x and $f(x)$ approach finite constants (a and A , respectively) as limits. It is also possible, however, for either one or both of x and $f(x)$ to become arbitrarily large or arbitrarily small. These types of limiting behavior are defined as follows:

If the difference between a function $f(x)$ and a constant A is less in absolute value than an arbitrarily small positive number for all positive values of x that are sufficiently large, then $f(x)$ is said to approach A as a limit when x becomes positively infinite, that is, increases without limit. This is indicated by the notation

$$\lim_{x \rightarrow \infty} f(x) = A \quad \text{or} \quad f(x) \rightarrow A \text{ as } x \rightarrow \infty$$

NOTE: The notation $\lim_{x \rightarrow +\infty} f(x) = A$ or $f(x) \rightarrow A$ as $x \rightarrow +\infty$ is sometimes used. However, ∞ is understood to mean $+\infty$, as would be the case if it were a number, and the plus sign is usually omitted for convenience.

EXAMPLE

If $f(x) = 1 - (1/x)$, then $\lim_{x \rightarrow \infty} f(x) = 1$, since, for example,

$$f(1) = 0$$

$$f(5) = \frac{4}{5}$$

$$f(20) = \frac{19}{20}$$

$$f(100) = \frac{99}{100}$$

$$f(1000) = \frac{999}{1000}$$

$$f(10,000) = \frac{9999}{10000}$$

etc.

Similarly, the limit of $f(x)$ may be defined when x becomes negatively infinite, that is, decreases without limit. This is indicated by the notation

$$\lim_{x \rightarrow -\infty} f(x) = A' \quad \text{or} \quad f(x) \rightarrow A' \text{ as } x \rightarrow -\infty$$

EXAMPLE

If $f(x) = 1 - (1/x)$, then

$$\lim_{x \rightarrow -\infty} f(x) = 1$$

If a function $f(x)$ is greater than an arbitrarily large positive number for all values of x that are sufficiently near a constant a and for which $x \neq a$, then $f(x)$ is said to become positively infinite (that is, increases without limit) as x approaches a . This is indicated by the notation

$$\lim_{x \rightarrow a} f(x) = \infty \quad \text{or} \quad f(x) \rightarrow \infty \text{ as } x \rightarrow a$$

Similarly, $f(x)$ becomes negatively infinite (that is, decreases without limit) when it assumes numerically large negative values. This is indicated by the notation

$$\lim_{x \rightarrow a} f(x) = -\infty \quad \text{or} \quad f(x) \rightarrow -\infty \text{ as } x \rightarrow a$$

EXAMPLE

If $f(x) = 1/(x-2)^2$, then

$$\lim_{x \rightarrow 2} f(x) = \infty$$

If a function $f(x)$ is greater than an arbitrarily large positive number for all positive values of x that are sufficiently large, then $f(x)$ is said to become positively infinite (that is, increases without limit) as x becomes positively infinite (that is, increases without limit). This is indicated by the notation

$$\lim_{x \rightarrow \infty} f(x) = \infty \quad \text{or} \quad f(x) \rightarrow \infty \text{ as } x \rightarrow \infty$$

The cases indicated by the following notation are defined similarly:

$$\lim_{x \rightarrow \infty} f(x) = -\infty \quad \text{or} \quad f(x) \rightarrow -\infty \text{ as } x \rightarrow \infty$$

$$\lim_{x \rightarrow -\infty} f(x) = \infty \quad \text{or} \quad f(x) \rightarrow \infty \text{ as } x \rightarrow -\infty$$

$$\lim_{x \rightarrow -\infty} f(x) = -\infty \quad \text{or} \quad f(x) \rightarrow -\infty \text{ as } x \rightarrow -\infty$$

EXAMPLE

If $f(x) = x^4 - 4$, then

$$\lim_{x \rightarrow \infty} f(x) = \infty \quad \text{and} \quad \lim_{x \rightarrow -\infty} f(x) = \infty$$

EXAMPLE

If $f(x) = x^3 - 8$, then

$$\lim_{x \rightarrow \infty} f(x) = \infty \quad \text{and} \quad \lim_{x \rightarrow -\infty} f(x) = -\infty$$

In some cases a function may approach either of two different limits, depending on whether the variable approaches its limit through values larger or smaller than that limit; in such a case, the limit is not defined (does not exist) but the right-hand and left-hand limits exist.

The *right-hand limit* of a function is the value the function approaches when the variable approaches its limit through decreasing values (that is, from the right); this type of limiting behavior is indicated by the notation

$$\lim_{x \rightarrow a^+} f(x) = A^+ \quad \text{or} \quad f(x) \rightarrow A^+ \text{ as } x \rightarrow a^+$$

The *left-hand limit* of a function is the value the function approaches when the variable approaches its limit through increasing values (that is, from the left); this is indicated by the notation

$$\lim_{x \rightarrow a^-} f(x) = A^- \quad \text{or} \quad f(x) \rightarrow A^- \text{ as } x \rightarrow a^-$$

Thus the limit of a function exists if and only if its right-hand and left-hand limits exist and are identical; in that case

$$\lim_{x \rightarrow a^+} f(x) = \lim_{x \rightarrow a^-} f(x) = \lim_{x \rightarrow a} f(x)$$

EXAMPLE

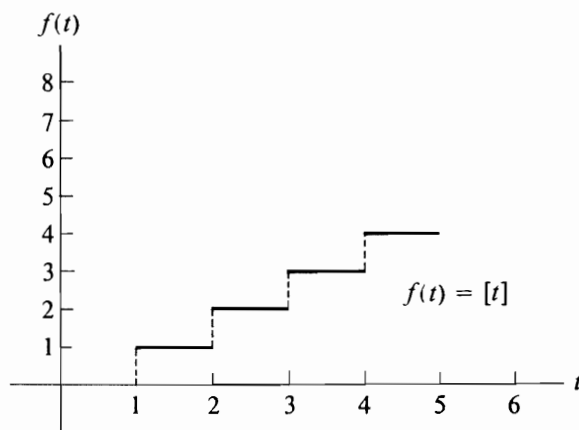
If $f(t) = [t] = \text{largest integer in } t$, then

$$\lim_{t \rightarrow 3^+} f(t) = 3 \quad \text{and} \quad \lim_{t \rightarrow 3^-} f(t) = 2$$

Thus $\lim_{t \rightarrow 3} f(t)$ is not defined.

It might seem that the limit of this function $f(t)$ as $t \rightarrow 3$ should be 3. However, when t is arbitrarily close to 3, some of the values of $[t]$ are 2 (when $t < 3$), others are 3 (when $t > 3$). Thus the values of t are not close to any one value A when t is arbitrarily close to 3 and $\lim_{t \rightarrow 3} [t]$ does not exist, although the right-hand and left-hand limits do exist. This is indicated in Figure 2.1 by the unbroken horizontal lines between successive integer values of t . That is, $f(t)$ approaches 2 as t approaches 3 from the left; $f(t)$ approaches 3 as t approaches 3 from the right. Clearly there is nothing unique about the integer 3; in fact, $f(t)$ does not have a limit (although it has right-hand and left-hand limits) as t approaches any integer value; $f(t)$ does, however, have the limit $[t]$ as t approaches any noninteger value.

Figure 2.1



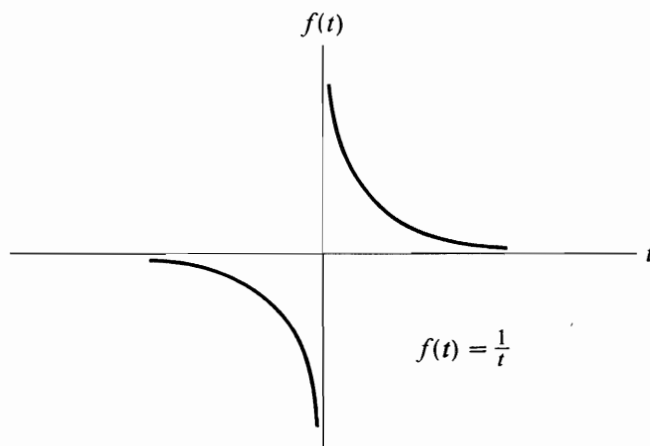
EXAMPLE

If $f(t) = 1/t$, then

$$\lim_{t \rightarrow 0^+} f(t) = \infty \quad \text{and} \quad \lim_{t \rightarrow 0^-} f(t) = -\infty$$

(see Figure 2.2)

Figure 2.2



Thus $\lim_{t \rightarrow 0} f(t)$ is not defined; as t approaches 0 from above, $f(t)$ becomes positively infinite; as t approaches 0 from below, $f(t)$ becomes negatively infinite.

It is also possible that even right-hand and left-hand limits do not exist; this is the case, for example, for the sine and cosine functions which are oscillatory in nature.

PROPERTIES OF LIMITS

The following properties of limits are useful in evaluating the limit of a function.

If $\lim_{x \rightarrow a} f(x) = b$, $\lim_{x \rightarrow a} g(x) = c$, and K is a constant,

$$1. \lim_{x \rightarrow a} K = K.$$

That is, the limit of a constant is equal to that constant. For purposes of interpreting the following properties of limits, a constant may be thought of as a special type of function—that is, a function that assumes only one value. Thus the following statements are valid (as special cases) if some or all of the functions are constants.

$$2. \lim_{x \rightarrow a} [f(x) \pm g(x)] = \lim_{x \rightarrow a} f(x) \pm \lim_{x \rightarrow a} g(x) = b \pm c.$$

That is, the limit of the sum (or difference) of two functions is equal to the sum (or difference) of their limits.

NOTE: This rule applies to the sum or difference of any finite number of functions, that is, if $\lim_{x \rightarrow a} f_i(x) = a_i$ for $i = 1, 2, \dots, n$, then

$$\lim_{x \rightarrow a} \left(\sum_{i=1}^n f_i(x) \right) = \sum_{i=1}^n \lim_{x \rightarrow a} f_i(x) = \sum_{i=1}^n a_i$$

$$3. \lim_{x \rightarrow a} [f(x) \cdot g(x)] = \left[\lim_{x \rightarrow a} f(x) \right] \left[\lim_{x \rightarrow a} g(x) \right] = bc.$$

That is, the limit of the product of two functions is equal to the product of their limits.

NOTE: This rule applies to the product of any finite number of functions, that is, if $\lim_{x \rightarrow a} f_i(x) = a_i$ for $i = 1, 2, \dots, n$, then

$$\lim_{x \rightarrow a} \left(\prod_{i=1}^n f_i(x) \right) = \prod_{i=1}^n \lim_{x \rightarrow a} f_i(x) = \prod_{i=1}^n a_i$$

$$4. \lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \frac{\lim_{x \rightarrow a} f(x)}{\lim_{x \rightarrow a} g(x)} = \frac{b}{c} \quad \text{if } c \neq 0.$$

That is, the limit of the quotient of two functions is equal to the quotient of their limits, provided the limit of the divisor is not zero.

$$5. \lim_{x \rightarrow a} [f(x)]^n = \left[\lim_{x \rightarrow a} f(x) \right]^n = b^n.$$

That is, the limit of the n th power of any function is equal to the n th power of the limit of the function. (This follows from rule 3 applied to the product of n equal factors.)

$$6. \lim_{x \rightarrow a} (\sqrt[n]{f(x)}) = \sqrt[n]{\lim_{x \rightarrow a} f(x)} = \sqrt[n]{b}.$$

That is, the limit of the principal n th root of a positive function is equal to the principal n th root of the limit of that function. (This follows from rule 4.)

EXAMPLE

$$\begin{aligned}\lim_{x \rightarrow 2} (3x + 5) &= \lim_{x \rightarrow 2} 3x + \lim_{x \rightarrow 2} 5 && \text{(rule 2)} \\ &= 6 + 5 && \text{(rules 3, 1)} \\ &= 11\end{aligned}$$

EXAMPLE

$$\begin{aligned}\lim_{x \rightarrow -1} \left(\frac{x^2}{x^2 + a} \right) &= \frac{\lim_{x \rightarrow -1} x^2}{\lim_{x \rightarrow -1} (x^2 + a)} && \text{(rule 4)} \\ &= \frac{\left(\lim_{x \rightarrow -1} x \right)^2}{\left(\lim_{x \rightarrow -1} x \right)^2 + \left(\lim_{x \rightarrow -1} a \right)} && \text{(rules 5, 2)} \\ &= \frac{1}{1 + a} && \text{(rule 1)}\end{aligned}$$

EXAMPLE

$$\begin{aligned}\lim_{x \rightarrow 5} (10\sqrt[3]{x^2 + 2}) &= \left(\lim_{x \rightarrow 5} 10 \right) \left[\lim_{x \rightarrow 5} (x^2 + 2) \right]^{1/3} && \text{(rules 3, 6)} \\ &= 10(27)^{1/3} && \text{(rules 1, 2, 5)} \\ &= 30\end{aligned}$$

EXAMPLE

$$\begin{aligned}\lim_{x \rightarrow 0} \left(\frac{x^2 + a}{x^2 + b} \right) &= \frac{\lim_{x \rightarrow 0} (x^2 + a)}{\lim_{x \rightarrow 0} (x^2 + b)} && \text{(rule 4)} \\ &= \frac{a}{b} && \text{(rules 2, 1, 5)}\end{aligned}$$

EXAMPLE

$$\begin{aligned}\lim_{x \rightarrow \infty} \left(\frac{a}{1 + \frac{1}{x}} \right) &= \frac{\lim_{x \rightarrow \infty} a}{\lim_{x \rightarrow \infty} \left(1 + \frac{1}{x} \right)} && \text{(rules 4, 2, 4)} \\ &= \frac{a}{1 + 0} && \text{(rule 1)} \\ &= a\end{aligned}$$

EXAMPLE

$$\begin{aligned}
 \lim_{x \rightarrow -\infty} \left(b - \frac{2}{3x + 10} \right) &= \lim_{x \rightarrow -\infty} b - \frac{\lim_{x \rightarrow -\infty} 2}{\lim_{x \rightarrow -\infty} (3x + 10)} && \text{(rules 2, 4)} \\
 &= b - 0 && \text{(rules 1, 2, 3)} \\
 &= b
 \end{aligned}$$

In applying rule 4 concerning the limit of a quotient, it may happen that the quotient of the limits takes the form $0/0$. This is one of the indeterminate forms discussed in some detail in a later section; however, it is frequently possible to obtain a determinate limit for the original quotient if a very simple procedure is used. This procedure consists of dividing both the numerator and the denominator by an expression which in the limit is equal to zero, thereby obtaining an expression having a limit.

EXAMPLE

If $f(x) = (x^2 - 4)/(x - 2)$, then $\lim_{x \rightarrow 2} f(x)$ is the indeterminate form $0/0$. But if $x \neq 2$, both numerator and denominator can be divided by $x - 2$. Then $f(x) = (x^2 - 4)/(x - 2) = x + 2$ and $\lim_{x \rightarrow 2} f(x) = 4$.

Recall that in evaluating $\lim_{x \rightarrow 2} f(x)$, it is not the value of $f(x)$ at $x = 2$ but only the values of $f(x)$ for x close to 2 which are of interest. It is assumed that x becomes arbitrarily close to 2, but does not equal 2. For $x \neq 2$, $f(x) = x + 2$ and $\lim_{x \rightarrow 2} f(x) = \lim_{x \rightarrow 2} (x + 2) = 4$. Thus, even though $f(x)$ is undefined at $x = 2$, it has the limit 4 as $x \rightarrow 2$.

EXAMPLE

If $F(h) = [(2 + h)^2 - 4]/h$, then $\lim_{h \rightarrow 0} F(h)$ is the indeterminate form $0/0$. However, for $h \neq 0$, $F(h) = (4 + 4h + h^2 - 4)/h = (4h + h^2)/h = 4 + h$ and thus $\lim_{h \rightarrow 0} F(h) = 4$.

A second type of indeterminate form that can arise as a result of applying rule 4 is ∞/∞ . This is another of the indeterminate forms discussed in later sections; in this case it is frequently possible to obtain a determinate limit for the original quotient if both the numerator and denominator are divided by the highest power of the variable which appears in the denominator.

EXAMPLE

If $f(x) = (2x^3 + x^2 - 3)/(x^3 + x + 2)$, then $\lim_{x \rightarrow \infty} f(x)$ is the indeterminate form ∞/∞ . However, if both the numerator and denominator are divided by x^3 , then

$$f(x) = \frac{2x^3 + x^2 - 3}{x^3 + x + 2} = \frac{2 + \frac{1}{x} - \frac{3}{x^3}}{1 + \frac{1}{x^2} + \frac{2}{x^3}}$$

and thus $\lim_{x \rightarrow \infty} f(x) = 2$, since as $x \rightarrow \infty$, $\frac{1}{x} \rightarrow 0$, $\frac{1}{x^2} \rightarrow 0$, $\frac{2}{x^3} \rightarrow 0$, and $\frac{3}{x^3} \rightarrow 0$.

EXAMPLE

If $f(x) = [x + (1/x)]/(2x - (1/x))$, then $\lim_{x \rightarrow \infty} f(x)$ is the indeterminate form ∞/∞ . However, if both the numerator and denominator are divided by x , then

$$f(x) = \frac{x + \frac{1}{x}}{2x - \frac{1}{x}} = \frac{1 + \frac{1}{x^2}}{2 - \frac{1}{x^2}}$$

and thus $\lim_{x \rightarrow \infty} f(x) = \frac{1}{2}$, since $1/x^2 \rightarrow 0$ as $x \rightarrow \infty$.

PROBLEMS

Evaluate the following limits.

1. $\lim_{t \rightarrow 2} (t^2 + 6t + 5)$
2. $\lim_{x \rightarrow 0} (x^2 + 6x + 3)$
3. $\lim_{y \rightarrow 5} \left(\frac{3y - 5}{y - 2} \right)$
4. $\lim_{y \rightarrow -2} (y - 3)^2$
5. $\lim_{x \rightarrow 2} (x^3 - 3x + 5)$
6. $\lim_{x \rightarrow 2} \left(\frac{x^2 \sqrt{x + 2}}{x^2 + 1} \right)$
7. $\lim_{x \rightarrow \infty} \left(\frac{2}{x^2 + 1} \right)$
8. $\lim_{x \rightarrow 0} \left(1 + \frac{1}{x^2} \right)$
9. $\lim_{x \rightarrow -1} \left(\frac{x^2 + 1}{x^2 + x + 1} \right)$
10. $\lim_{x \rightarrow \infty} \left(4 - \frac{2}{x + 1} \right)$
11. $\lim_{t \rightarrow 0} \left(\frac{3t - 5}{t + 2} \right)$
12. $\lim_{t \rightarrow \infty} \left(\frac{3t^2 - 5t + 4}{t^2 + 2} \right)$
13. $\lim_{h \rightarrow 0} a^{x+h}$
14. $\lim_{h \rightarrow 0} 2^{-h}$
15. $\lim_{x \rightarrow 2} x^{-4}$
16. $\lim_{x \rightarrow -2} x^4$
17. $\lim_{x \rightarrow -\infty} 2^x$
18. $\lim_{x \rightarrow \infty} 2^{-x}$
19. $\lim_{t \rightarrow \infty} e^{-t}$
20. $\lim_{t \rightarrow 0} \left(\frac{e^t + e^{-t}}{2} \right)$
21. $\lim_{x \rightarrow 0} \left(\frac{x - y}{x + y} \right)$
22. $\lim_{x \rightarrow 0} \left(\frac{2^x - 2^{-x}}{2^x + 2^{-x}} \right)$
23. $\lim_{y \rightarrow 0} \left(\frac{x - y}{x + y} \right)$
24. $\lim_{x \rightarrow \infty} \left(\frac{x^3 - 2x + 5}{2x^3 - 7} \right)$
25. $\lim_{x \rightarrow \infty} \left(\frac{x^2 + a^2}{x^3 + a^3} \right)$
26. $\lim_{x \rightarrow 0} \left(\frac{x^3 - 5x + 6}{x^2 - 2x + 3} \right)$
27. $\lim_{x \rightarrow \infty} \left(\frac{1}{1 + 2^{1/x}} \right)$
28. $\lim_{x \rightarrow -\infty} \left(\frac{1}{1 + 2^{1/x}} \right)$
29. $\lim_{x \rightarrow 0^+} \left(\frac{1}{1 + 2^{1/x}} \right)$
30. $\lim_{x \rightarrow 0^-} \left(\frac{1}{1 + 2^{1/x}} \right)$
31. $\lim_{x \rightarrow -2} x^3$
32. $\lim_{x \rightarrow 2} x^{-3}$
33. $\lim_{x \rightarrow a} \left(\frac{ax + 10}{x} \right)$
34. $\lim_{x \rightarrow \infty} \left(\frac{1}{1 + e^{1/x}} \right)$
35. $\lim_{x \rightarrow 0^-} \left(\frac{1}{1 + e^{1/x}} \right)$

36. $\lim_{x \rightarrow 0^+} \left(\frac{1}{1 + e^{1/x}} \right)$
37. $\lim_{x \rightarrow 0} \left(\frac{x + e^{x^2+3} + x^2}{e^3 + x} \right)$
38. $\lim_{x \rightarrow 0} \left(\frac{x}{1 + e^x} \right)$
39. $\lim_{h \rightarrow 1} \left(\frac{e^{-h}}{h^3 + 4h + 5} \right)$
40. $\lim_{t \rightarrow 2} \left[\frac{t^2 + 4}{(t+2)(t+3)} \right]$
41. $\lim_{t \rightarrow \infty} (e^{1/t} + 5)$
42. $\lim_{t \rightarrow 0} \left(\frac{e^t - e^{-2t} - e^{3t^2}}{10} \right)$
43. $\lim_{x \rightarrow 4} \left[\frac{x^2 - 16}{(x-4)^2} \right]$
44. $\lim_{x \rightarrow -4} \left[\frac{x^2 - 16}{(x-4)^2} \right]$
45. $\lim_{x \rightarrow 4} \left[\frac{x^2 - 16}{(x+4)^2} \right]$
46. $\lim_{x \rightarrow 0} \left(\frac{h^{-x} + h^x}{x} \right)$
47. $\lim_{t \rightarrow 2} e^{-t+2}$
48. $\lim_{h \rightarrow 0} \left(\frac{e^{-h} - e^{2h}}{3} \right)$
49. $\lim_{x \rightarrow \infty} (1 + 3^{1/x})$
50. $\lim_{y \rightarrow 0} \left(\frac{x - y + 3}{x + y - 6} \right)$
51. $\lim_{x \rightarrow 3} \left(\frac{x^3 - 3x^2 + 2x - 6}{x + 4} \right)$
52. $\lim_{h \rightarrow 0} \left[\frac{(h+1)e^{-h}}{h^2 + 1} \right]$
53. $\lim_{x \rightarrow \infty} e^{1/x}$
54. $\lim_{y \rightarrow 0} \left(\frac{x^2 - y^2 + 2}{x^3 - y} \right)$
55. $\lim_{x \rightarrow 2} \left(\frac{x^4 - 6x - 4}{x + 1} \right)$
56. $\lim_{t \rightarrow \infty} \left(\frac{t^3 + 4t^2 + 10}{5t^2 + 12t} \right)$
57. $\lim_{h \rightarrow \infty} \left(\frac{1 + e^{1/h}}{e^h} \right)$
58. $\lim_{h \rightarrow -\infty} \left(\frac{1 + e^{1/h}}{e^h} \right)$
59. $\lim_{h \rightarrow 2} \left(\frac{x - h}{x + h} \right)$
60. $\lim_{t \rightarrow 2} \left(\frac{t^2 - 6t + 8}{t^2 - 5t + 6} \right)$
61. $\lim_{h \rightarrow \infty} \left(\frac{h^4 + 5h^5}{3h + 2h^6} \right)$
62. $\lim_{t \rightarrow 0} \left(\frac{e^{-t}}{1 + e^{1/t}} \right)$
63. $\lim_{t \rightarrow -3} \left(\frac{t^2 - t - 12}{t^2 + 4t + 3} \right)$
64. $\lim_{x \rightarrow 0} \left(\frac{x^2 + 2ax + a^2}{a^3} \right)$
65. $\lim_{x \rightarrow 0^+} \left(\frac{2 - e^x}{e^{1/x}} \right)$
66. $\lim_{x \rightarrow 0^-} \left(\frac{2 - e^x}{e^{1/x}} \right)$

Answers to Odd-Numbered Problems

1. 21
3. $\frac{10}{3}$
5. 7
7. 0
9. 2
11. $-\frac{5}{2}$
13. a^x
15. $\frac{1}{16}$
17. 0
19. 0
21. -1
23. 1
25. 0
27. $\frac{1}{2}$
29. 0
31. -8
33. $\frac{a^2 + 10}{a}$
35. 1
37. 1
39. $\frac{1}{10e}$

41. 6
 43. ∞
 45. 0
 47. 1
 49. 2
 51. 0
 53. 1
 55. 0

57. 0

59. $\frac{x-2}{x+2}$

61. 0

63. $\frac{7}{2}$

65. 0

2.3 Continuity

In this section continuity of a function is defined. Essentially, a function is continuous if its graph consists of an unbroken curve; the mathematical definition of continuity involves properties of limits. Types of discontinuities are discussed first and then properties of continuous functions are given.

In the definition of $\lim_{x \rightarrow a} f(x)$ the value of $f(x)$ for $x = a$ is not specified; that is, this limit depends only on the values of $f(x)$ in the neighborhood of (i.e., close to) $x = a$ but not on the value of $f(x)$ at $x = a$. Thus $\lim_{x \rightarrow a} f(x)$ may or may not be equal to $f(a)$. If $\lim_{x \rightarrow a} f(x)$ exists and the value $f(a)$ exists and is equal to $\lim_{x \rightarrow a} f(x)$, then $f(x)$ is continuous at $x = a$. That is, a function $f(x)$ is said to be *continuous at $x = a$* if

1. $f(a)$ is defined.
2. $\lim_{x \rightarrow a} f(x)$ exists.
3. $\lim_{x \rightarrow a} f(x) = f(a)$.

NOTE: When a limit is said to exist, this should be understood to mean that the limit exists finitely. Possible confusion arises because, for example, the expression $\lim_{x \rightarrow a} A = \infty$ is written; however, this does not mean that A approaches a number designated by ∞ , but only that A becomes arbitrarily large as x approaches a . It should be remembered that ∞ is *not* a number and should not be thought of as a number, even though for convenience it is used as a number would be used in some expressions.

A function $f(x)$ is said to be *continuous in (or on) an interval $b \leq x \leq c$* (or $b < x < c$) if it is continuous at every point of the interval.

A function that is not continuous at a point $x = a$ is said to be *discontinuous at $x = a$* .

From the definition of continuity, it follows that the graph of a function that is continuous on an interval consists of an unbroken curve (that is, a curve that can be drawn without raising the pen from the paper) over that interval. Continuity was tacitly assumed for the curves discussed in the preceding sections concerning graphical representation. In subsequent sections this assumption, which makes it possible to sketch a curve by plotting relatively few points and drawing an unbroken curve through them, will be justified for several (large) classes of curves.

TYPES OF DISCONTINUITIES

The three conditions that must be satisfied for a function to be continuous can be violated in various ways, resulting in different types of discontinuities. In general, three types of discontinuities occur.

1. A function $f(x)$ is said to have an *infinite discontinuity* at $x = a$ if $f(x)$ becomes infinite (positively or negatively) as $x \rightarrow a$. That is, $f(a)$ is not defined and $\lim_{x \rightarrow a} f(x)$ does not exist.
2. A function $f(x)$ is said to have a *finite discontinuity* at $x = a$ if $f(x)$ remains finite but changes abruptly at $x = a$. That is, $f(a)$ is defined but $\lim_{x \rightarrow a} f(x)$ does not exist [although, in general, the right-hand and left-hand limits exist and $f(a)$ is equal to one of them].
3. A function $f(x)$ is said to have a *missing-point discontinuity* at $x = a$ if $f(a)$ is not defined but $\lim_{x \rightarrow a} f(x)$ exists.

It should be noted that a *discrete function* is defined only for a finite number of values of x in any interval and is thus discontinuous at all other (infinitely many) values of x in the interval. Since x assumes only discrete values, the concept of a limit is not appropriate for discrete functions.

Geometrically, the curves representing functions having different types of discontinuities are quite different in appearance: The curve representing a function having an infinite discontinuity at $x = a$ approaches $x = a$ as an asymptote; the curve representing a function having a finite discontinuity at $x = a$ has an abrupt jump or step at $x = a$; the curve representing a function having a missing-point discontinuity at $x = a$ appears to be continuous, but the single point $x = a$ is missing; the curve representing a discrete function consists of discrete points and thus is not a curve in the usual sense.

A missing-point discontinuity at $x = a$ occurs because $f(a)$ is not defined; this type of discontinuity may be removed by defining

$$f(a) = \lim_{x \rightarrow a} f(x)$$

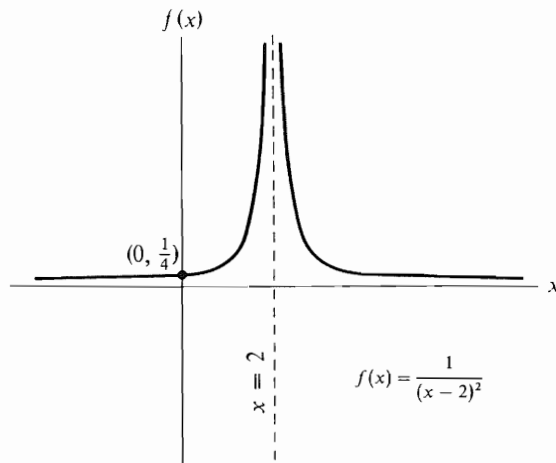
since all three conditions of the definition of a continuous function at $x = a$ are then satisfied. It is possible for a function to have more than one missing-point discontinuity; as long as the number of these discontinuities is finite, they may be removed by the procedure described. Note that there is also the logical possibility (although a practical example is difficult to find) that there is a discontinuity at $x = a$ because $f(a)$, although defined, is not equal to $\lim_{x \rightarrow a} f(x)$. Such a discontinuity is not removable, since $f(a)$ is already defined and cannot therefore be defined as equal to $\lim_{x \rightarrow a} f(x)$.

Although missing-point discontinuities are removable, finite and infinite discontinuities are not removable. This is the case because the definition of a function is an arbitrary set of rules; adding another rule to this set, provided it is compatible with the others, is perfectly permissible, and this is all that is required to remove a missing-point discontinuity. On the other hand, finite and infinite discontinuities occur because the function in question does not approach a limit under the specified conditions. A limit either exists or does not exist as a mathematical fact associated with the nature of the function in question and cannot be defined into existence; thus there is no way to remove a discontinuity resulting from the nonexistence of a limit.

EXAMPLE

The function $f(x) = 1/(x - 2)^2$ has an infinite discontinuity at $x = 2$, since $f(x) \rightarrow \infty$ as $x \rightarrow 2$ and $f(2)$ is undefined. However, this function is continuous at all values of x other than $x = 2$ (see Figure 2.3).

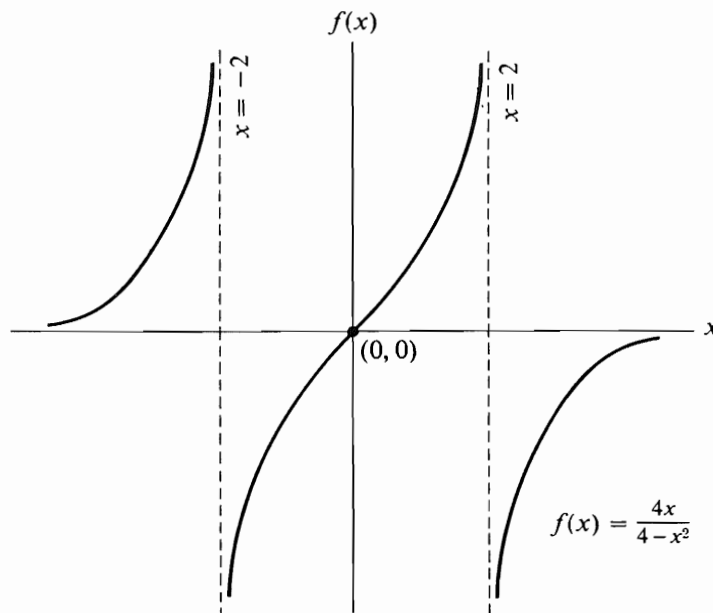
Figure 2.3



EXAMPLE

The function $f(x) = 4x/(4 - x^2)$ has infinite discontinuities at $x = \pm 2$, since $f(x) \rightarrow \infty$ as $x \rightarrow 2^-$, $f(x) \rightarrow -\infty$ as $x \rightarrow 2^+$, $f(x) \rightarrow \infty$ as $x \rightarrow -2^-$, and $f(x) \rightarrow -\infty$ as $x \rightarrow -2^+$, and $f(+2)$ and $f(-2)$ are undefined. However, this function is continuous for all values of x other than $x = \pm 2$ (see Figure 2.4).

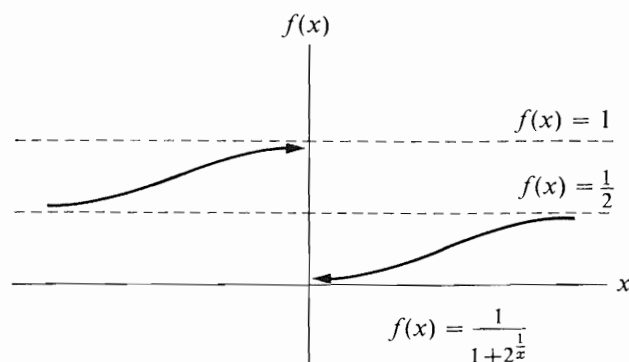
Figure 2.4



EXAMPLE

The function $f(x) = 1/(1 + 2^{1/x})$ has a finite discontinuity (that is, a “jump”) at $x = 0$, since $\lim_{x \rightarrow 0} f(x)$ is undefined. However, this function is continuous at all values of x other than $x = 0$ (see Figure 2.5).

Figure 2.5



EXAMPLE

The function $f(x) = (x^3 - 2x^2 - 3x + 6)/(x - 2)$ is undefined and is thus discontinuous for $x = 2$. However, for $x \neq 2$,

$$f(x) = \frac{x^3 - 2x^2 - 3x + 6}{x - 2} = \frac{(x^2 - 3)(x - 2)}{x - 2} = x^2 - 3$$

and $\lim_{x \rightarrow 2} f(x) = 1$. Thus if, by definition,

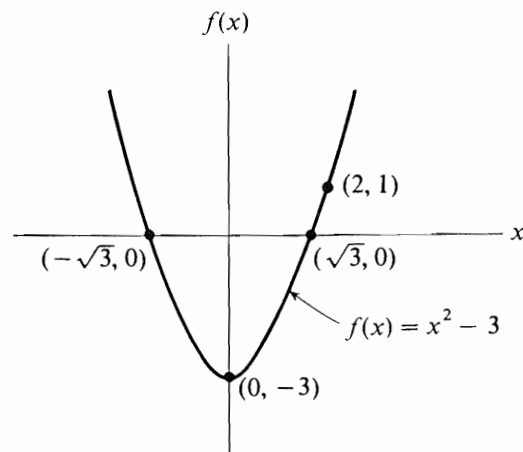
$$f(2) = \lim_{x \rightarrow 2} f(x) = 1$$

then $f(x)$ is continuous for all x and its graph is the parabola

$$y = x^2 - 3$$

(see Figure 2.6).

Figure 2.6



EXAMPLE

The function $f(x) = (x^2 - 9)/(x + 3)$ is undefined and is thus discontinuous for $x = -3$. However, for $x \neq -3$,

$$f(x) = \frac{x^2 - 9}{x + 3} = x - 3$$

and $\lim_{x \rightarrow -3} f(x) = -6$. Thus if, by definition,

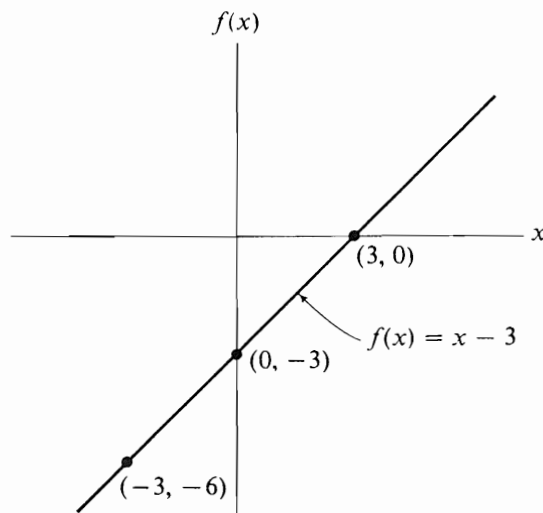
$$f(-3) = \lim_{x \rightarrow -3} f(x) = -6$$

then $f(x)$ is continuous for all x and its graph is the straight line

$$y = x - 3$$

(see Figure 2.7).

Figure 2.7



Many functions in business and economics are discrete or have finite discontinuities of the step-function type. For example, price and cost functions are frequently discrete because of the nature of the commodity involved or have discontinuities because cost and price per unit decrease (or increase) abruptly for particular quantities. Supply and demand functions and many other economic functions are also frequently discrete because of the nature of the commodity involved. It should be noted that functions which are, in fact, discrete are frequently represented as continuous for convenience; this applies, for example, to the supply and demand functions for commodities sold in units—such as refrigerators, eggs, light bulbs, chairs, lawn mowers, cars, and so forth. Representation as continuous of a function which is by nature discrete makes possible the use of many tools of analysis not otherwise applicable. However, in interpreting the results of such analyses, the inherent discreteness should not be forgotten—for example, it is inappropriate to discuss the price of 1.632 refrigerators or the wages of 29.2 workers.

EXAMPLE

A wholesale grocer sells number 2 size cans of a certain vegetable in case lots according to the following price schedule:

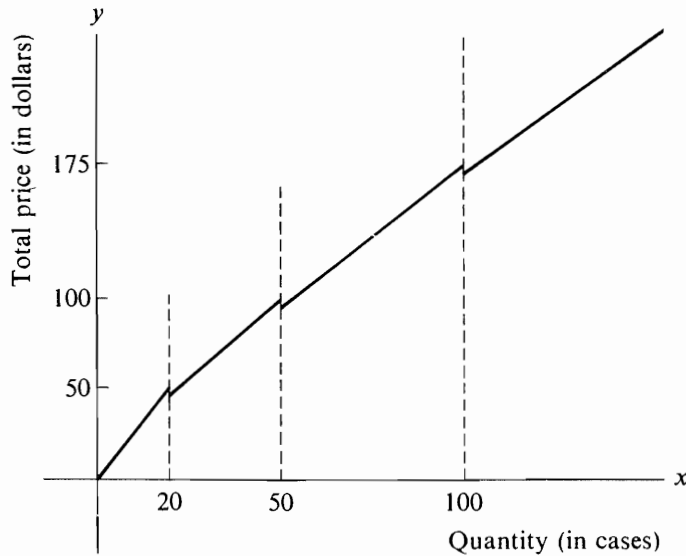
- \$2.50 per case for 20 cases or less
- \$2.00 per case for orders of more than 20 cases and not more than 50 cases
- \$1.75 per case for orders of more than 50 cases and not more than 100 cases
- \$1.50 per case for orders of more than 100 cases

If y is total price and x is quantity in cases, the price function can be represented algebraically as follows.

$$y = \begin{cases} 2.50x & 0 \leq x \leq 20 \\ 2.00x & 20 < x \leq 50 \\ 1.75x & 50 < x \leq 100 \\ 1.50x & x > 100 \end{cases}$$

and geometrically as shown in Figure 2.8.

Figure 2.8



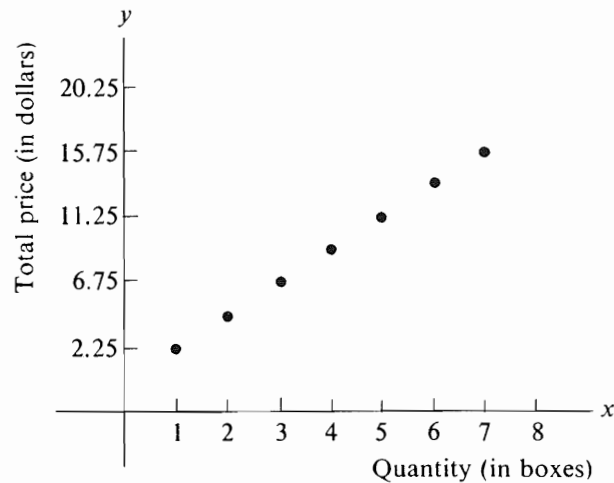
Note that, for this price function, total price is not a monotonic function of quantity ordered. This makes certain orders unlikely. For example, the price of 17 cases is \$42.50, while the price of 21 cases is only \$42.00; thus an order of 17 cases would probably be unlikely.

EXAMPLE

A company sells printed business stationery only in boxes of 200 sheets at \$2.25 per box. If y is total price and x is number of boxes, the price function can be represented algebraically by the equation

$$y = 2.25x \quad \text{for} \quad x = 1, 2, 3, \dots$$

Figure 2.9

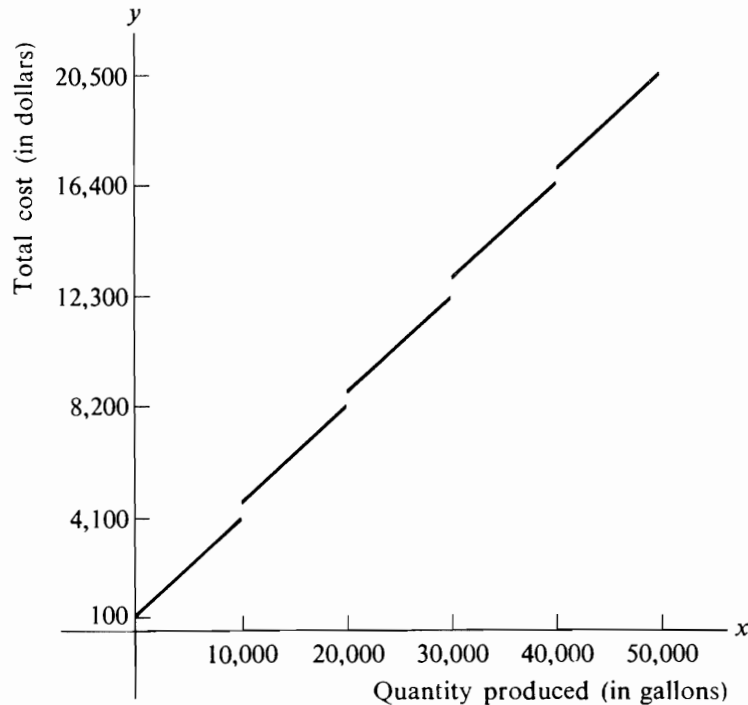


and geometrically by the graph in Figure 2.9. This function is discrete, since it is defined only for integer values of x .

EXAMPLE

An oil refinery has 5 distillation towers and operates as many of them as needed to process the raw materials available. The overhead cost of operating each distillation tower (operator, maintenance, and so forth) is \$100 per week; in addition, the cost of raw materials is \$0.40 per

Figure 2.10



gallon of refined oil. Each distillation tower can process raw materials yielding 10,000 gallons of refined oil per week. If y is the total cost of operation and x is the quantity of refined oil (in gallons), the cost function can be represented algebraically by the equation

$$f(x) = 100 \left(\left\lfloor \frac{x}{10,000} \right\rfloor + 1 \right) + 0.4x$$

where $\lfloor x/10,000 \rfloor$ denotes the largest integer part of $x/10,000$. Note that 1 is added to $\lfloor x/10,000 \rfloor$ because $\lfloor x/10,000 \rfloor$ distillation towers produce a maximum of $10,000x$ gallons. Thus for any additional production, another distillation tower must be operated. The graph of the function is shown in Figure 2.10. This function has finite discontinuities at $x = 10,000, 20,000, 30,000,$ and $40,000$.

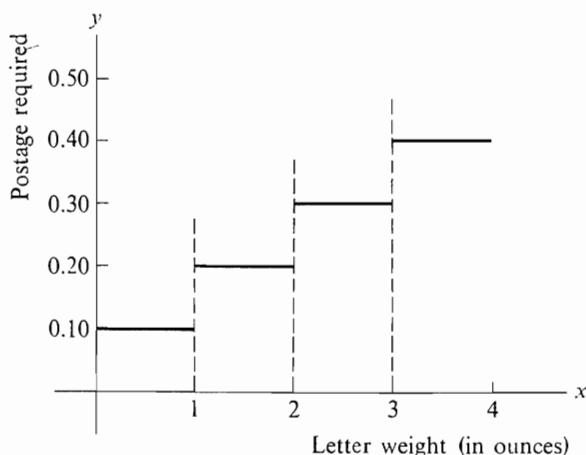
EXAMPLE

United States first-class postage is 10 cents per ounce or fraction of an ounce. If y is postage required and x is letter weight (in ounces), the cost function can be represented by the equation

$$y = 0.10(\lfloor x \rfloor + 1)$$

where $\lfloor x \rfloor$ denotes the largest integer part of x . The graph of the function is shown in Figure 2.11. This function has finite discontinuities at all integer values of x .

Figure 2.11



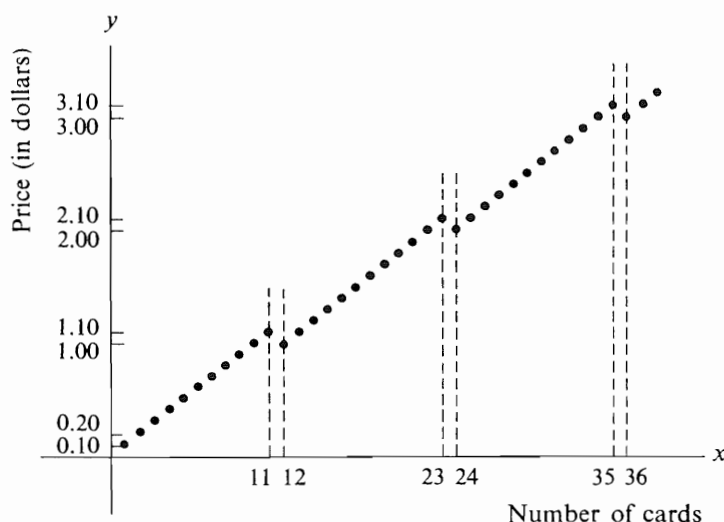
EXAMPLE

A stationery store sells Christmas cards for 10 cents each or \$1.00 for a dozen. If y is price and x is number of cards sold, the price function for a customer can be represented by the equation

$$y = \left\lfloor \frac{x}{12} \right\rfloor + 0.10 \left(x - 12 \left\lfloor \frac{x}{12} \right\rfloor \right) \quad \text{for } x = 1, 2, 3, \dots$$

where $\lfloor x/12 \rfloor$ denotes the largest integer part of $x/12$. This function is discrete since it is defined only for integer values of x ; it is also nonmonotonic, as a result of discounting for larger quantities purchased. (See Figure 2.12.)

Figure 2.12



PROPERTIES OF CONTINUOUS FUNCTIONS

The following properties of continuous functions follow from the definition of a continuous function and the properties of limits.

If $f(x)$ and $g(x)$ are two functions that are continuous at $x = a$, then

$$F(x) = f(x) + g(x)$$

$$G(x) = f(x) - g(x)$$

$$H(x) = f(x) \cdot g(x)$$

are also continuous at $x = a$. If $g(a) \neq 0$, then

$$I(x) = \frac{f(x)}{g(x)}$$

is also continuous at $x = a$. [Note that if $g(a) = 0$, then $I(a)$ is undefined.]

The following properties can also be proved (but by more advanced arguments):

If a function $f(x)$ is continuous on a closed interval $a \leq x \leq b$, then

1. $f(x)$ has a greatest value (maximum) and a least value (minimum) in the interval.
2. If k is a number between a and b , then there is at least one point c between a and b such that $f(c) = k$.
3. If $f(a)$ and $f(b)$ are opposite in sign, then there exists at least one point d in the interval such that $f(d) = 0$.

PROBLEMS

For each of the following functions, determine the values of x for which the function is discontinuous, identify the types of discontinuities, and give appropriate definitions for removing any removable discontinuities.

$$1. f(x) = \frac{3x + 5}{x^2 + 4x + 4}$$

$$2. f(x) = \frac{x - 2}{x + 1}$$

3. $f(x) = \frac{x^2 - 1}{x - 1}$

4. $f(x) = \frac{x^2 + 4}{x^2 - x - 2}$

5. $f(x) = \log(2x - 5)$

6. $f(x) = \frac{1}{2^x - 1}$

7. $f(x) = \frac{1}{x(x - 2)}$

8. $f(x) = \sqrt{\frac{x+1}{x-1}}$

9. $f(x) = \frac{(x-1)^2}{(x-1)(x^2 - 4x + 5)}$

10. $f(x) = \log\left(\frac{x-2}{x}\right)$

11. $f(x) = \frac{x^2 + 5x + 6}{x + 2}$

12. $f(x) = \frac{x^2 - 3}{x^2 - 4}$

13. $f(x) = \frac{x^2 - 2x}{x^3 - x^2 + x}$

14. $f(x) = \frac{x^2 - 5x + 6}{x - 2}$

15. $f(x) = \frac{x^2 + 1}{x^3 - 4x}$

16. $f(x) = \frac{4x}{4 - x^2}$

17. $f(x) = \frac{1}{4x^2 - 16}$

18. $f(x) = \frac{x - 2}{(x - 2)(x^2 + 2x + 10)}$

19. $f(x) = \frac{1}{e^{4x} - 1}$

20. $f(x) = \frac{x - 2}{(x - 2)(x^2 + 2x - 3)}$

21. $f(x) = \frac{e^x + 2x^2}{2e^{3x} - 2}$

22. $f(x) = \frac{x^2 - 5x + 6}{(x - 2)(x^2 - 3x + 5)}$

23. $f(x) = \frac{1}{e^x - 1}$

24. $f(x) = \frac{x^2 - 3}{x^2 - 16}$

25. $f(x) = \frac{(x - 3)^3}{(x - 3)(x^2 - 2x + 6)}$

26. $f(x) = \frac{x^2 + 2x - 8}{x + 4}$

27. $f(x) = \frac{1}{3e^{3x} - 3}$

28. $f(x) = \frac{(x + 5)^2(x + 3)}{(x + 5)(x^2 - 4x + 8)}$

29. $f(x) = \frac{x^2 + 5x + 6}{x + 2}$

30. $f(x) = \frac{x^2 + 3x + 6}{x^2 - 1}$

31. $f(x) = \frac{10}{e^{6x} - 1}$

32. $f(x) = \frac{x^2 + x - 2}{x^2 + 27x + 50}$

33. $f(x) = \frac{x^2 + 10x + 1}{x^2 - 9}$

34. $f(x) = \frac{x^2 - 5x + 4}{x - 4}$

35. $f(x) = \ln(x^2 - 6)$

36. $f(x) = \frac{x^2 - 5x + 6}{(x - 2)(x^2 - 2x - 3)}$

37. $f(x) = \frac{2x^2 + 3x}{x^3 - 9x}$

38. $f(x) = \frac{1}{3e^x - 3}$

39. $f(x) = \frac{e^x + 4x}{3e^{4x} - 3}$

40. $f(x) = \frac{x^2 + 5x + 4}{(x + 4)(x^2 - 6x + 10)}$

41. $f(x) = \frac{x^2 - 1}{x + 1}$ for $x \neq -1$

$$= -2 \quad \text{for } x = -1$$

Is $f(x)$ continuous at $x = -1$?

42. $f(x) = \frac{x^2 + 3x - 10}{x - 2}$ for $x \neq 2$.

What value should be assigned to $f(2)$ to make $f(x)$ continuous at $x = 2$?

43. $f(x) = \sqrt[3]{x}$. Is $f(x)$ continuous at $x = 0$?

44. $f(x) = \frac{x^2 - 7x + 12}{x - 3}$ for $x \neq 3$

$$= 1 \quad \text{for } x = 3$$

Is $f(x)$ continuous at $x = 3$?

Answers to Odd-Numbered Problems

- | | |
|-------------------------|---|
| 1. $x = -2$ | 23. $x = 0$ |
| 3. $x = 1, f(1) = 2$ | 25. $x = 3, f(3) = 0$ |
| 5. $x \leq \frac{5}{2}$ | 27. $x = 0$ |
| 7. $x = 0, x = 2$ | 29. $x = -2, f(-2) = 1$ |
| 9. $x = 1, f(1) = 0$ | 31. $x = 0$ |
| 11. $x = -2, f(-2) = 1$ | 33. $x = \pm 3$ |
| 13. $x = 0, f(0) = -2$ | 35. $x = \pm \sqrt{6}$ |
| 15. $x = 0, x = \pm 2$ | 37. $x = \pm 3, x = 0, f(0) = -\frac{1}{3}$ |
| 17. $x = \pm 2$ | 39. $x = 0$ |
| 19. $x = 0$ | 41. continuous |
| 21. $x = 0$ | 43. continuous |

2.4 Definition of the First Derivative

In this section the first derivative of a function is defined and several interpretations of the first derivative are considered. The first derivative of a function at a point is the slope of the function at that point. This concept is defined precisely as follows:

The slope m of a straight line is defined as the tangent of its angle of inclination or, equivalently, as the ratio of the change in vertical distance (rise) to the change in horizontal distance (run) as a point moves along the line in either direction (see Figure 2.13).

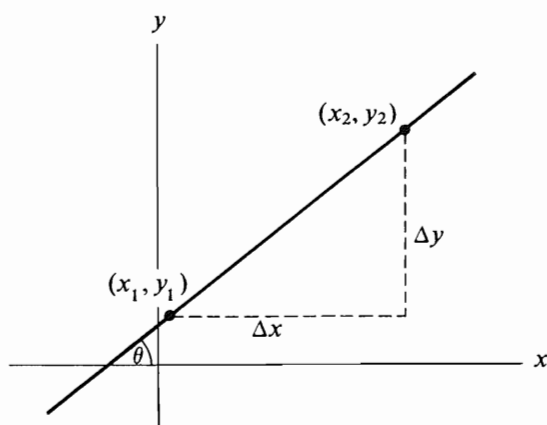
$$m = \tan \theta = \frac{y_2 - y_1}{x_2 - x_1} = \frac{\Delta y}{\Delta x}$$

The slope of any given straight line is a constant—that is, the rate of change of y as x changes is constant over the length of the line. However, for other curves the slope is not constant and must be determined for each particular point of interest.

Suppose that (x_1, y_1) and (x_2, y_2) are any two points on the curve $y = f(x)$. Then the slope of the line (called a secant) joining (x_1, y_1) and (x_2, y_2) is

$$m_{\text{sec}} = \frac{y_2 - y_1}{x_2 - x_1} = \frac{\Delta y}{\Delta x}$$

Figure 2.13



Suppose now that the point (x_1, y_1) is held fixed while the point (x_2, y_2) is moved along the curve $y = f(x)$ toward the point (x_1, y_1) ; as the point (x_2, y_2) is moved along the curve $f(x)$, the slope of the line joining (x_1, y_1) and (x_2, y_2) will, in general, vary. However, it may happen, and does happen for most curves encountered in practice, that as the point (x_2, y_2) moves closer and closer to the point (x_1, y_1) , the slope of the secant line varies by smaller and smaller amounts and, in fact, approaches a constant limiting value. When this happens, the limiting value is said to be the slope of the tangent to the curve at (x_1, y_1) or the slope of the curve at (x_1, y_1) [see Figure 2.14].

More concisely, if, as the point (x_2, y_2) approaches the point (x_1, y_1) along the curve $y = f(x)$, the slope of the secant line approaches a constant limiting value, then this limiting value is said to be the slope of the tangent to the curve at (x_1, y_1) or, briefly, the slope of the curve at (x_1, y_1) , or the slope of $f(x)$ at (x_1, y_1) . That is,

$$\lim_{\Delta x \rightarrow 0} m_{\text{sec}} = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \text{slope of } f(x) \text{ at } (x_1, y_1)$$

The slope of a function at a given point is the *first derivative* of the function at the given point.

It is customary in defining the first derivative to use the notation (x, y) for the stationary point (x_1, y_1) and the notation $(x + \Delta x, y + \Delta y)$ for the moving point (x_2, y_2) . Then

$$\frac{dy}{dx} = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

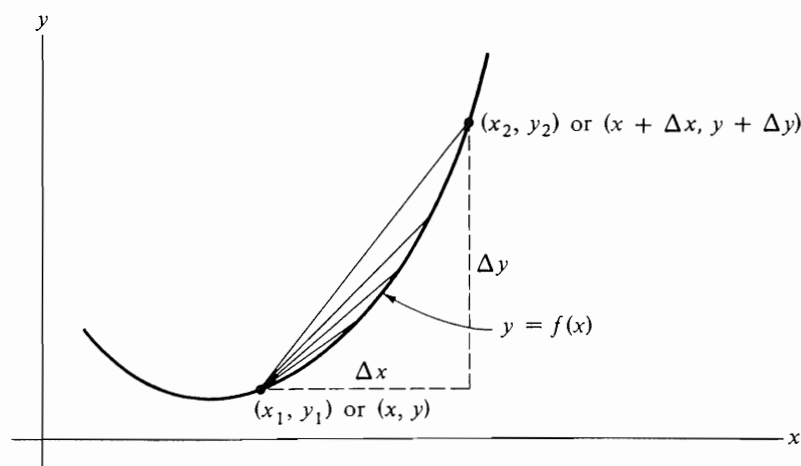
is the first derivative with respect to x of the function $y = f(x)$. This limit may exist for some values of x and fail to exist for other values of x . At each point (x, y) where this limit does exist, the function $y = f(x)$ is said to have a derivative or to be

differentiable and $\frac{dy}{dx}$ is said to be the *first derivative* or the derivative of $y = f(x)$. The process of obtaining the first derivative of a function is referred to as *differentiation*.

Various types of notation, in addition to $\frac{dy}{dx}$, are used to denote the first derivative of $y = f(x)$ with respect to x . The most common of these are

$$f'(x) \quad y' \quad \frac{d}{dx}(y) \quad D_x y \quad D_x(y)$$

Figure 2.14



It should be noted that the first derivative of a function with respect to x is, in general, another function of x which must be evaluated for particular values of interest. This corresponds to the fact stated above that, except for straight lines, curves do not in general have the same slope at different points.

From the definition of a derivative and the properties of limits, rules for obtaining the derivatives of various types of functions can be obtained. These are given in a later section. For illustration, the derivatives of several relatively simple functions are obtained directly rather than by using the formulas.

The steps for obtaining a derivative directly are as follows:

1. Consider the function $y = f(x)$.
2. Give increments to x and y to obtain

$$y + \Delta y = f(x + \Delta x)$$

3. Subtract to obtain

$$\Delta y = f(x + \Delta x) - f(x)$$

4. Divide by Δx to obtain

$$\frac{\Delta y}{\Delta x} = \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

5. Take the limit to obtain

$$\frac{dy}{dx} = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

EXAMPLE

Find the first derivative of $y = 4x + 1$.

$$\begin{aligned} \frac{dy}{dx} &= \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \\ &= \lim_{\Delta x \rightarrow 0} \frac{4(x + \Delta x) + 1 - 4x - 1}{\Delta x} \\ &= \lim_{\Delta x \rightarrow 0} \frac{4\Delta x}{\Delta x} \\ &= \lim_{\Delta x \rightarrow 0} 4 = 4 \end{aligned}$$

Note that $y = 4x + 1$ represents a straight line and thus $\frac{dy}{dx}$ is a constant.

EXAMPLE

Find the first derivative of $y = x^3 - 12x + 13$.

$$\begin{aligned} \frac{dy}{dx} &= \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \\ &= \lim_{\Delta x \rightarrow 0} \frac{(x + \Delta x)^3 - 12(x + \Delta x) + 13 - x^3 + 12x - 13}{\Delta x} \\ &= \lim_{\Delta x \rightarrow 0} \frac{x^3 + 3x(\Delta x)^2 + 3x^2(\Delta x) + (\Delta x)^3 - 12(x + \Delta x) + 13 - x^3 + 12x - 13}{\Delta x} \end{aligned}$$

$$\begin{aligned}
& [\text{NOTE: } (a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3] \\
& = \lim_{\Delta x \rightarrow 0} \frac{3x^2(\Delta x) + 3x(\Delta x)^2 + (\Delta x)^3 - 12(\Delta x)}{\Delta x} \\
& = \lim_{\Delta x \rightarrow 0} [3x^2 + 3x(\Delta x) + (\Delta x)^2 - 12] \\
& = 3x^2 - 12
\end{aligned}$$

Note that $\frac{dy}{dx}$ is a function of x and can be evaluated for any value of x .

EXAMPLE

Find the first derivative of $y = x^2 + (1/x)$ for $x \neq 0$.

$$\begin{aligned}
\frac{dy}{dx} &= \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \\
&= \lim_{\Delta x \rightarrow 0} \frac{(x + \Delta x)^2 + \frac{1}{x + \Delta x} - x^2 - \frac{1}{x}}{\Delta x} \\
&= \lim_{\Delta x \rightarrow 0} \frac{x^2 + 2x(\Delta x) + (\Delta x)^2 + \frac{1}{x + \Delta x} - x^2 - \frac{1}{x}}{\Delta x} \\
&= \lim_{\Delta x \rightarrow 0} \frac{2x(\Delta x) + (\Delta x)^2 + \frac{x - x - \Delta x}{x(x + \Delta x)}}{\Delta x} \\
&= \lim_{\Delta x \rightarrow 0} \left[2x + \Delta x - \frac{1}{x(x + \Delta x)} \right] \\
&= 2x - \frac{1}{x^2} \quad \text{for } x \neq 0
\end{aligned}$$

Note that again $\frac{dy}{dx}$ is a function of x and can be evaluated for any value of $x \neq 0$.

PROBLEMS

For each of the following functions $f(x)$, find the first derivative $\frac{dy}{dx}$ using the definition

$$\frac{dy}{dx} = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

1. $f(x) = x^2 - x + 1$

4. $f(x) = x - \frac{1}{x}$

2. $f(x) = \frac{1}{x^2}$

5. $f(x) = 6 - 2x^3$

3. $f(x) = \sqrt{2x}$

6. $f(x) = 5x^4 - 2$

Answers to Odd-Numbered Problems

1. $2x - 1$

5. $-6x^2$

3. $\frac{1}{\sqrt{2x}}$

INTERPRETATION OF THE FIRST DERIVATIVE

Although the geometrical interpretation of the first derivative has historical significance and is perhaps the most intuitive, other interpretations are equally valid and sometimes are more useful in applications.

Two of the classical interpretations of a derivative are in terms of the velocity of a moving body and the rate of change of a function. Although both of these interpretations originated in the study of various problems in physics and mathematics, they have since been applied in many fields. For example, marginal analysis in economics can be more readily understood in terms of the rate of change of a function.

Velocity of a Moving Body

Consider a body (or a particle) moving along a straight-line path. Let t denote the time measured from some fixed instant and let s denote the distance of the particle from some fixed origin on the line, where s is positive or negative according to the direction of movement from the origin. Suppose that distance from the origin is given in terms of time by a function $s = f(t)$, called the *law of motion*. [For example, think of the body or particle as being a car that is driven along a straight road in such a way that its distance from the starting point is given as a function of t by $s = f(t)$.]

At a certain time t_1 , let the particle be at a distance s_1 from the origin 0, and suppose that during the following time interval Δt it moves a distance Δs farther from the origin (see Figure 2.15).

If the ratio $\frac{\Delta s}{\Delta t}$ is constant, so that equal distances are always traversed in equal intervals of time, the motion is said to be *uniform* and the ratio $\frac{\Delta s}{\Delta t}$ is called the *velocity* at any instant. The term *speed* is used frequently to mean the magnitude—that is, absolute value—of the velocity.

If, however, the motion is not uniform, the ratio $\frac{\Delta s}{\Delta t}$ varies as Δt varies and is no longer the velocity of the particle at any instant. It is instead the *average velocity* of the particle during the particular time interval Δt :

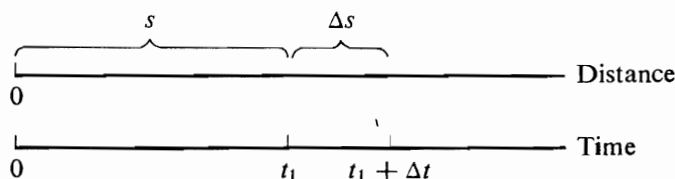
$$\text{average velocity during interval } \Delta t = \frac{\Delta s}{\Delta t}$$

As the time interval Δt approaches zero, this average velocity $\frac{\Delta s}{\Delta t}$ may approach a limit. If so, this limit is said to be the *instantaneous velocity* at time t_1 :

$$\text{instantaneous velocity at time } t_1 = \lim_{\Delta t \rightarrow 0} \frac{\Delta s}{\Delta t}$$

But, by definition, $\lim_{\Delta t \rightarrow 0} \frac{\Delta s}{\Delta t}$ is the first derivative of $f(t)$ at the point $t = t_1$. Thus, at an instant t_1 , the velocity of a particle moving in a straight line according to the

Figure 2.15



law of motion $s = f(t)$, where s is the directed distance from a fixed origin and t is the time, is given by the value of the derivative of s with respect to t for $t = t_1$.

EXAMPLE

The distance of a train from its starting point when it is traveling along a straight track is given by the equation

$$s = 16t^2 + 2t$$

where s is distance in miles and t is time in hours. Find the distance traveled and the velocity after 2 hours.

$$s = 16t^2 + 2t$$

So after 2 hours, $s = 64 + 4 = 68$ miles traveled.

$$\begin{aligned} \frac{ds}{dt} &= \lim_{\Delta t \rightarrow 0} \frac{f(t + \Delta t) - f(t)}{\Delta t} \\ &= \lim_{\Delta t \rightarrow 0} \frac{16(t + \Delta t)^2 + 2(t + \Delta t) - 16t^2 - 2t}{\Delta t} \\ &= \lim_{\Delta t \rightarrow 0} \frac{16t^2 + 32t(\Delta t) + 16(\Delta t)^2 + 2t + 2\Delta t - 16t^2 - 2t}{\Delta t} \\ &= \lim_{\Delta t \rightarrow 0} \frac{32t(\Delta t) + 16(\Delta t)^2 + 2(\Delta t)}{\Delta t} \\ &= \lim_{\Delta t \rightarrow 0} [32t + 16(\Delta t) + 2] \\ &= \lim_{\Delta t \rightarrow 0} (32t + 2) \end{aligned}$$

and the velocity after 2 hours is

$$\left. \frac{ds}{dt} \right|_{t=2} = 64 + 2 = 66 \text{ miles per hour}$$

(The notation $\left. \frac{ds}{dt} \right|_{t=2}$ means the function $\frac{ds}{dt}$ evaluated for $t = 2$.)

EXAMPLE

One of the attractions at a carnival is a “prove your strength” bell, which rings if a lever is hit with sufficient strength to propel an iron ball up a vertical shaft to hit the bell. When the lever is hit with a force of 50 lb, the distance of the ball from the bottom of the shaft is given by

$$s = 40t - 32t^2$$

where s is distance in feet and t is time in minutes since the lever was hit.

- If the bell is at the top of a 15-ft shaft, is a force of 50 lb sufficient to ring it?
- How far from the bottom of the shaft is the ball after 30 seconds, 45 seconds, 1 minute, and 75 seconds, respectively?

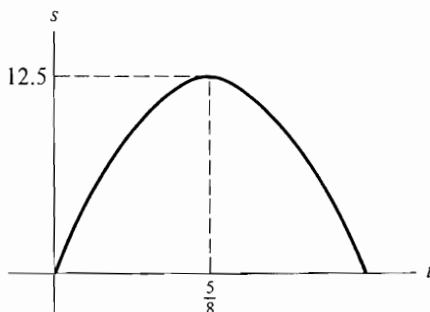
(a) Velocity $\frac{ds}{dt}$ is zero when the ball is at the top of the shaft, so solving the equation $\frac{ds}{dt} = 0$ provides the time the ball travels up the shaft before coming back down.

$$\begin{aligned}
\frac{ds}{dt} &= \lim_{\Delta t \rightarrow 0} \frac{f(t + \Delta t) - f(t)}{\Delta t} \\
&= \lim_{\Delta t \rightarrow 0} \frac{40(t + \Delta t) - 32(t + \Delta t)^2 - 40t + 32t^2}{\Delta t} \\
&= \lim_{\Delta t \rightarrow 0} \frac{40t + 40(\Delta t) - 32t^2 - 64t(\Delta t) - 32(\Delta t)^2 - 40t + 32t^2}{\Delta t} \\
&= \lim_{\Delta t \rightarrow 0} \frac{40(\Delta t) - 64t(\Delta t) - 32(\Delta t)^2}{\Delta t} \\
&= \lim_{\Delta t \rightarrow 0} [40 - 64t - 32(\Delta t)] \\
&= 40 - 64t
\end{aligned}$$

If $\frac{ds}{dt} = 0$, $t = \frac{40}{64} = \frac{5}{8}$ minutes (or 37.5 seconds) and $s = 40(\frac{5}{8}) - 32(\frac{25}{64}) = 12.5$ feet, so the force is not sufficient to ring the bell 15 ft up the shaft. (See Figure 2.16.)

- (b) If $t = 30$ seconds $= \frac{1}{2}$, $s = 40(\frac{1}{2}) - 32(\frac{1}{4}) = 12$ ft (on the way up)
 If $t = 45$ seconds $= \frac{3}{4}$, $s = 40(\frac{3}{4}) - 32(\frac{9}{16}) = 12$ ft (on the way down)
 If $t = 1$, $s = 40 - 32 = 8$ ft (on the way down)
 If $t = 75$ seconds $= \frac{5}{4}$, $s = 40(\frac{5}{4}) - 32(\frac{25}{16}) = 0$ (back down)

Figure 2.16



Rate of Change of a Function

Let p and q denote the measures of two related variables and regard q as a function of p —that is, $q = f(p)$. If the ratio $\frac{\Delta q}{\Delta p}$ of corresponding changes in the two quantities has the same value for all values of Δp , it is called the *rate of change* of q with respect to p and q is said to change uniformly with respect to p . But if the ratio $\frac{\Delta q}{\Delta p}$ is not constant as Δp changes, q does not change uniformly, and $\frac{\Delta q}{\Delta p}$ is then called the *average rate of change* of q with respect to p over the interval Δp .

If the ratio $\frac{\Delta q}{\Delta p}$ approaches a limit when Δp approaches zero, then this limit is said to be the *instantaneous rate of change* of q with respect to p .

From the definition of the derivative, it follows that the instantaneous rate of change of a variable quantity q with respect to a related variable quantity p is given by the derivative of q with respect to p , that is, $\frac{dq}{dp}$.

It is customary to use the expression “rate of change of a function” as equivalent to the derivative of the function.

If the variable q can be expressed as a function of the variable time t , then the derivative $\frac{dq}{dt}$ of q with respect to t gives the *time rate of change* of q . Thus the velocity of a particle is the time rate of change of its distance from the origin.

The interpretation of a derivative in terms of the rate of change of a function is frequently applicable in economics, as discussed below.

2.5 Applications of the First Derivative in Business and Economics

In this section several applications of the concept of the first derivative in business and economics are discussed. These applications include marginal cost, marginal revenue, elasticity, and marginal propensity to save and to consume.

It is customary in economics to describe the variation of one quantity y with respect to another quantity x in terms of two concepts, *average* and *marginal*. The average concept expresses the variation of y over a range of values of x , usually the range from zero up to a certain selected value. The marginal concept, on the other hand, concerns the variation of y “on the margin”—that is, for very small variations of x from a given value. The marginal concept is thus precise only when it is considered in the limiting sense, as the variation in x approaches zero.

The economic concepts of average and marginal variation correspond to the more general concepts of the average rate of change of a function over an interval and the instantaneous rate of change (that is, the derivative) of a function.

Average and marginal variations in quantities are essential considerations in the development of both microeconomic and macroeconomic theory. Examples of the application of the derivative in both microeconomic theory (cost, revenue, elasticity) and macroeconomic theory (income, consumption, savings) are discussed below; additional examples are given in later sections.

COST

Suppose the total cost y of producing and marketing x units of a commodity is given by the function

$$y = f(x)$$

Then the *average cost* per unit is

$$\frac{y}{x} = \frac{f(x)}{x}$$

If output is increased by an amount Δx from a certain level x and if the corresponding increase in cost is Δy , then the average increase in cost per unit increase in output is $\frac{\Delta y}{\Delta x}$ and marginal cost is defined as

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \frac{dy}{dx} = f'(x)$$

That is, marginal cost is the derivative $F'(x)$ of the total cost function $y = f(x)$ with respect to x and is the rate of increase in total cost with increase in output.

EXAMPLE

Consider the total cost function

$$y = 20 + 2x + 0.5x^2$$

where y denotes total cost and x denotes quantity produced. Average cost is

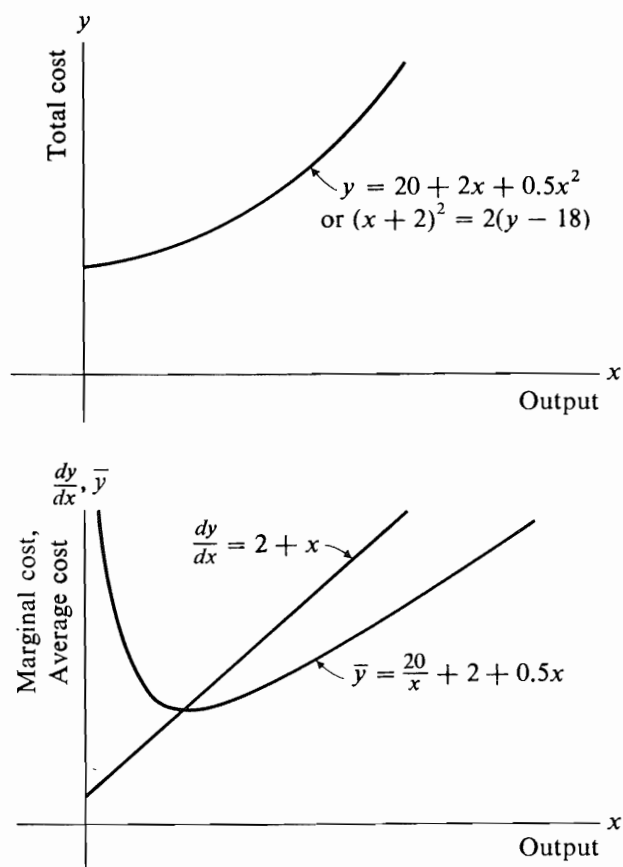
$$\bar{y} = \frac{y}{x} = \frac{20}{x} + 2 + 0.5x$$

and marginal cost is

$$\frac{dy}{dx} = 2 + x$$

Thus total cost increases as output increases; average cost per unit decreases and then increases as output increases; and marginal cost (rate of increase in total cost) increases as output increases (see Figure 2.17).

Figure 2.17



NOTE: Derivatives are given in these examples without any explanation of how they were obtained. The method of the preceding section or the rules given in the next section can be used to verify the derivations.

REVENUE

For any given demand function

$$y = f(x)$$

where y is the price per unit and x is the number of units, *total revenue* R is the product of x and y —that is,

$$R = xy = x \cdot f(x)$$

Marginal revenue with respect to demand is the derivative with respect to x of total revenue,

$$\frac{dR}{dx} = R'(x)$$

and is thus the rate of change in revenue with change in demand.

Note that *average revenue*, or revenue per unit, also represents the price per unit, y —that is, the average revenue curve and the demand curve are identical.

Since x and y are always nonnegative within our previously stated analytical framework, R is also always nonnegative. However, $\frac{dR}{dx}$ may be positive or negative—that is, although total revenue is always nonnegative, it may increase or decrease as demand increases.

EXAMPLE

Consider the demand function

$$3x + 4y = 10$$

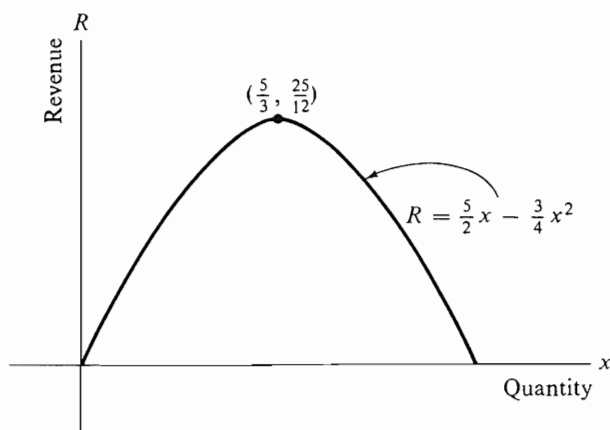
where y is the price per unit and x is the number of units. Then

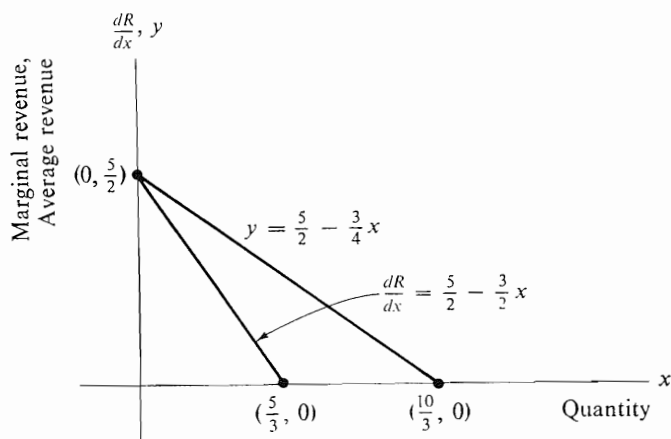
$$y = \frac{5}{2} - \frac{3}{4}x$$

Total revenue is

$$R = \frac{5}{2}x - \frac{3}{4}x^2$$

Figure 2.18





and marginal revenue is

$$\frac{dR}{dx} = \frac{5}{2} - \frac{3}{2}x$$

(see Figure 2.18).

For the example above, note the following properties, which hold for all linear demand functions. Total revenue increases and then decreases as quantity increases; both average revenue and marginal revenue decrease (linearly) as quantity increases. The average and marginal revenue curves have the same y -intercept and, apart from this common point, the marginal revenue curve lies below the average revenue curve. The marginal revenue curve intersects the x -axis at the output for which total revenue is greatest and the average revenue curve intersects the x -axis at twice that output. The slope of the average revenue curve is one-half the slope of the marginal revenue curve.

ELASTICITY

The *point elasticity* of the function $y = f(x)$ at the point x is the proportional rate of change in y per unit change in x :

$$\frac{Ey}{Ex} = \frac{\frac{dy}{y}}{\frac{dx}{x}} = \frac{x}{y} \cdot \frac{dy}{dx}$$

Note that the elasticity of a function is independent of the units in which the variables are measured. This results from the definition of elasticity in terms of proportional changes, which are necessarily independent of units of measurement.

The point elasticity of demand, supply, cost, productivity, and other functions is an important concept in economic theory. Detailed examples of these and other applications of derivatives are discussed in later sections, following the discussion of the techniques of differentiation.

NATIONAL INCOME, CONSUMPTION, AND SAVINGS

The relationship between total national disposable income and total national consumption is often referred to as the *consumption function*. In simple forms of theoretical analysis of the consumption function, it is assumed that as income increases (or decreases) consumption increases (or decreases), but by less than the increase (or decrease) in income; that is, the marginal propensity to consume is greater than zero but less than 1, where marginal propensity to consume is the rate of change in consumption as disposable income changes.

If the consumption function is given by

$$c = f(x)$$

where c is total national consumption and x is total national income (and c and x are in the same units), then the marginal propensity to consume is

$$\frac{dc}{dx} = f'(x)$$

In elementary theoretical analyses of national income, the assumption is often made that disposable income equals consumption plus savings. This is expressed as

$$x = c + s$$

The *marginal propensity to consume* is

$$\frac{dc}{dx} = f'(x)$$

and the *marginal propensity to save* is

$$\frac{ds}{dx} = 1 - \frac{dc}{dx}$$

In national income analysis, investment is considered as capital formation and represents an increase in real capital as typified by equipment, buildings, inventories, and so forth. Investment and consumption are assumed to be related in such a way that an initial investment expenditure may result in an increase in income of several times that amount. A precise numerical expression for this relationship is given by the multiplier. The multiplier is the ratio of the ultimate increase in income to the increase in investment that gave rise to it.

The *multiplier* k is related to the marginal propensity to consume and is given by

$$k = \frac{1}{1 - \frac{dc}{dx}} = \frac{1}{\frac{ds}{dx}}$$

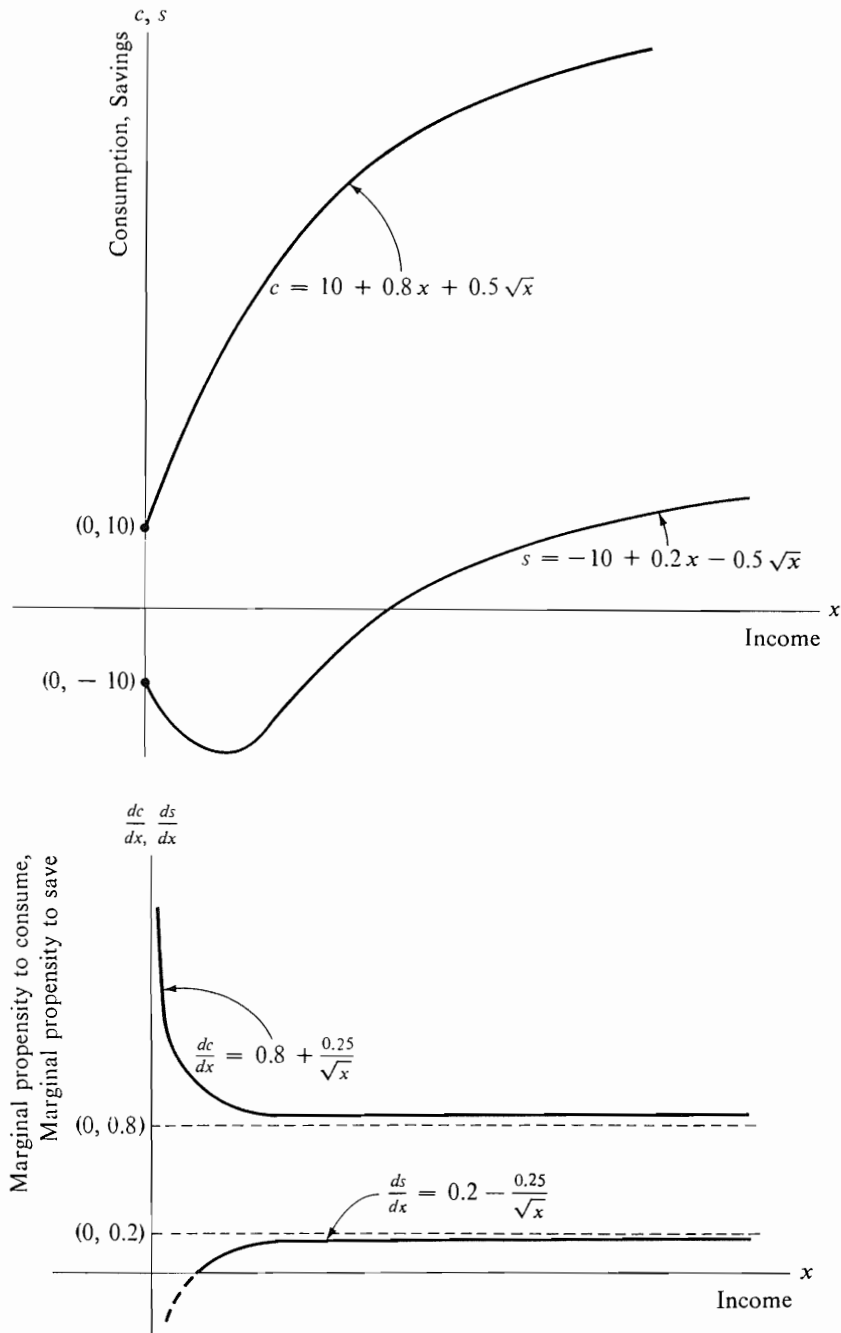
Note that if $\frac{dc}{dx} = 0$, then $k = 1$; that is, if none of the additional income is spent, the total increase in income is equal to the initial expenditure; if $\frac{dc}{dx} \rightarrow 1$, $k \rightarrow \infty$; that is, if all the additional income is spent, the total increase in income becomes infinitely large.

EXAMPLE

Suppose the consumption function is

$$c = 10 + 0.8x + 0.5\sqrt{x}$$

Figure 2.19



where c is total consumption and x is total disposable income. Then the amount saved is

$$\begin{aligned}
 s &= x - c \\
 &= x - (10 + 0.8x + 0.5\sqrt{x}) \\
 &= -10 + 0.2x - 0.5\sqrt{x}
 \end{aligned}$$

Note that c is positive for all positive values of x ; s is negative for small positive values of x and positive for larger values of x ; for every value of x , $c + s = x$. Note also that c is an increasing

function of x ; s is at first a decreasing and then an increasing function of x . (See Figure 2.19.)

The marginal propensity to consume is

$$\frac{dc}{dx} = 0.8 + \frac{0.25}{\sqrt{x}}$$

The marginal propensity to save is

$$\frac{ds}{dx} = 0.2 - \frac{0.25}{\sqrt{x}}$$

and

$$\frac{dc}{dx} + \frac{ds}{dx} = 1. \text{ The multiplier is}$$

$$k = \frac{1}{0.2 - \frac{0.25}{\sqrt{x}}}$$

Note that $\frac{dc}{dx}$ and k are decreasing functions of x and $\frac{ds}{dx}$ is an increasing function of x . (See Figure 2.19.)

NOTE: For certain economic analyses, it is more appropriate to consider income as a function of consumption rather than consumption as a function of income. In this case, income is plotted on the y -axis and consumption is plotted on the x -axis.

2.6 Rules for Differentiation

The procedure for obtaining derivatives directly was illustrated for several simple functions in previous sections. Although straightforward, this procedure is tedious and likely to result in errors, especially for more complicated functions. In practice, derivatives are obtained using rules or formulas for differentiating particular types of functions. These rules can be derived using the general procedure, as shown in Technical Note I.

In this section, rules for differentiating polynomial, algebraic, logarithmic, exponential, trigonometric, inverse, and composite functions are stated and illustrated.

ALGEBRAIC FUNCTIONS

The formulas in this section apply to the products, quotients, powers, and roots of differentiable functions.

Rule 1. The derivative of a constant is zero: If $y = c$,

$$\frac{dy}{dx} = 0$$

EXAMPLE

If $y = 6$, $\frac{dy}{dx} = 0$.

EXAMPLE

If $y = a$, $\frac{dy}{dx} = 0$.

Rule 2. The derivative of the n th power of a variable is the product of n and the $(n - 1)$ st power of the variable: If $y = x^n$,

$$\frac{dy}{dx} = nx^{n-1}$$

for any positive integer n .

EXAMPLE

If $y = x^2$, $\frac{dy}{dx} = 2x$.

EXAMPLE

If $y = x^{1/4}$, $\frac{dy}{dx} = \frac{1}{4}x^{-3/4}$.

EXAMPLE

If $y = x^{-4}$, $\frac{dy}{dx} = -4x^{-5} = -\frac{4}{x^5}$.

EXAMPLE

If $y = x^{-5/3}$, $\frac{dy}{dx} = -\frac{5}{3}x^{-8/3} = -\frac{5}{3x^{8/3}}$.

Rule 3. The derivative of the product of a constant and a differentiable function is the product of the constant and the derivative of the function: If $y = cu$, where $u = f(x)$ is a differentiable function of x ,

$$\frac{dy}{dx} = c \frac{du}{dx}$$

EXAMPLE

If $y = 10x$, $\frac{dy}{dx} = 10$.

EXAMPLE

$$\text{If } y = 3x^2, \quad \frac{dy}{dx} = 6x.$$

EXAMPLE

$$\text{If } y = -2x^{4/3}, \quad \frac{dy}{dx} = -\left(\frac{4}{3}\right)(2)x^{1/3} = -\frac{8}{3}x^{1/3}.$$

EXAMPLE

$$\text{If } y = -3x^{-1}, \quad \frac{dy}{dx} = -(-1)(3)x^{-2} = 3x^{-2} = \frac{3}{x^2}.$$

Rule 4. The derivative of the sum of a finite number of differentiable functions is the sum of their derivatives: If $y = u + v$, where $u = f(x)$ and $v = g(x)$ are differentiable functions of x ,

$$\frac{dy}{dx} = \frac{du}{dx} + \frac{dv}{dx}$$

In general, if $y = \sum_{i=1}^n u_i$, where $u_i = f_i(x)$ are differentiable functions of x for $i = 1, 2, \dots, n$, then

$$\frac{dy}{dx} = \sum_{i=1}^n \frac{du_i}{dx}$$

EXAMPLE

$$\text{If } y = 3x^2 + 4x + 2, \quad \frac{dy}{dx} = 6x + 4$$

EXAMPLE

$$\text{If } y = 5x^{1/2} + 3x^{1/4} + 7, \quad \frac{dy}{dx} = \frac{5}{2}x^{-1/2} + \frac{3}{4}x^{-3/4}$$

EXAMPLE

$$\text{If } y = 10 - 6x^{-1/2}, \quad \frac{dy}{dx} = -\left(-\frac{1}{2}\right)(6)x^{-3/2} = 3x^{-3/2}$$

EXAMPLE

$$\text{If } y = 6x^5 + x^4 + 2x^{3/2} + 5, \quad \frac{dy}{dx} = 30x^4 + 4x^3 + 3x^{1/2}$$

Rule for Differentiation of Polynomial Functions

These four rules for differentiation can be summarized and generalized as the rule for differentiation of polynomial functions. A single term of the form cx^n , where c is a constant and n is zero or a positive integer, is called a *monomial* in x . A function that is the sum of a finite number of such monomial terms is called a *polynomial* or a *polynomial function* in x .

The derivative of a monomial in x is

$$\frac{d}{dx}(cx^n) = cnx^{n-1}$$

and the derivative of the sum of a finite number of such terms (that is, the derivative of a polynomial) is the sum of the derivatives of the terms:

$$\frac{d}{dx}\left(\sum_{i=1}^n c_i x^{n_i}\right) = \sum_{i=1}^n c_i n_i x^{n_i-1}$$

PROBLEMS

Find the first derivative with respect to x for each of the following functions $y = f(x)$.

- | | |
|-----------------------------|------------------------------------|
| 1. $y = 6x + 2$ | 7. $y = 5x^{3/5} + 6x$ |
| 2. $y = x^3 + x$ | 8. $y = 7x^2 + 8x^{-1/2} + 2$ |
| 3. $y = 4x^2 + 2x$ | 9. $y = 9x^3 + 5x^{-1/4}$ |
| 4. $y = x^{1/2} + 4$ | 10. $y = 10x^4 + 2x^{5/4}$ |
| 5. $y = 3x^2 - x^{1/3} + 2$ | 11. $y = 12x^{3/2} + 6x^{1/2} + 2$ |
| 6. $y = 5x^{-1/3} + 5$ | 12. $y = x^5 + 3x^{-7/3} + 5$ |

Answers to Odd-Numbered Problems

- | | |
|-------------------------------|----------------------------------|
| 1. 6 | 7. $3x^{-2/5} + 6$ |
| 3. $8x + 2$ | 9. $27x^2 - \frac{5}{4}x^{-5/4}$ |
| 5. $6x - \frac{1}{3}x^{-2/3}$ | 11. $18x^{1/2} + 3x^{-1/2}$ |

Rule 5. The derivative of the product of two differentiable functions is equal to the first function times the derivative of the second function plus the second function times the derivative of the first function. Similarly, the product of more than two differentiable functions is the sum of the products of the derivative of each function and the other functions. If $y = uv$, where $u = f(x)$ and $v = g(x)$ are differentiable functions of x ,

$$\frac{dy}{dx} = u \frac{dv}{dx} + v \frac{du}{dx}$$

In general, if $y = \prod_{i=1}^n u_i$, where $u_i = f_i(x)$ are differentiable functions of x for $i = 1, 2, \dots, n$, then

$$\frac{d}{dx} \left(\prod_{i=1}^n u_i \right) = \sum_{i=1}^n \left[f'_i(x) \prod_{\substack{j=1 \\ j \neq i}}^n u_j \right]$$

EXAMPLE

If $y = (x^3 + 4)(x + 3)$,

$$\frac{dy}{dx}(uv) = u \frac{dv}{dx} + v \frac{du}{dx}$$

and

$$\frac{du}{dx} = 3x^2$$

$$\frac{dv}{dx} = 1(x^{1-1}) = 1$$

$$\frac{dy}{dx} = (x^3 + 4)(1) + (x + 3)(3x^2)$$

$$= x^3 + 4 + 3x^3 + 9x^2$$

$$= 4x^3 + 9x^2 + 4$$

EXAMPLE

If $y = (\sqrt{x} + 3)(x^2 + 6)$,

$$\frac{dy}{dx}(uv) = u \frac{dv}{dx} + v \frac{du}{dx}$$

and

$$\frac{du}{dx} = \frac{1}{2}x^{-1/2}$$

$$\frac{dv}{dx} = 2x$$

$$\frac{dy}{dx} = (\sqrt{x} + 3)(2x) + (x^2 + 6)\left(\frac{1}{2}x^{-1/2}\right)$$

$$= 2x^{3/2} + 6x + \frac{1}{2}x^{3/2} + 3x^{-1/2}$$

$$= \frac{5}{2}x^{3/2} + 6x + 3x^{-1/2}$$

EXAMPLE

If $y = (3x + 7)(x^{-2} + 8)$,

$$\frac{d}{dx}(3x + 1) = 3$$

$$\frac{d}{dx}(x^{-2} + 8) = -2x^{-3}$$

and

$$\begin{aligned}\frac{dy}{dx} &= (3x + 7)(-2x^{-3}) + (x^{-2} + 8)(3) \\ &= -6x^{-2} - 14x^{-3} + 3x^{-2} + 24 \\ &= -3x^{-2} - 14x^{-3} + 24\end{aligned}$$

EXAMPLE

If $y = (x + 3)(2x + 3)(x^2 + 1)$,

$$\frac{d}{dx}(x + 3) = 1$$

$$\frac{d}{dx}(2x + 3) = 2$$

$$\frac{d}{dx}(x^2 + 1) = 2x$$

and

$$\begin{aligned}\frac{dy}{dx} &= (1)(2x + 3)(x^2 + 1) + (2)(x + 3)(x^2 + 1) + (2x)(x + 3)(2x + 3) \\ &= (4x + 9)(x^2 + 1) + 2x(2x^2 + 9x + 9) \\ &= 4x^3 + 9x^2 + 4x + 9 + 4x^3 + 18x^2 + 18x \\ &= 8x^3 + 27x^2 + 22x + 9\end{aligned}$$

Rule 6. The derivative of the quotient of two differentiable functions is the quotient of the product of the denominator and the derivative of the numerator minus the product of the numerator and the derivative of the denominator divided by the square of the denominator: If $y = u/v$, where $u = f(x)$ and $v = g(x)$ are differentiable functions of x ,

$$\frac{dy}{dx} = \frac{v \frac{du}{dx} - u \frac{dv}{dx}}{v^2}$$

EXAMPLE

If $y = \frac{x^2 - 4x + 1}{x - 6}$,

$$u = x^2 - 4x + 1 \quad v = x - 6$$

$$\frac{du}{dx} = 2x - 4 \quad \frac{dv}{dx} = 1$$

and

$$\begin{aligned}
 \frac{dy}{dx} &= \frac{(x-6)(2x-4) - (x^2-4x+1)(1)}{(x-6)^2} \\
 &= \frac{2x^2 - 16x + 24 - x^2 + 4x - 1}{(x-6)^2} \\
 &= \frac{x^2 - 12x + 23}{(x-6)^2}
 \end{aligned}$$

EXAMPLE

If $y = \frac{4}{x^6}$,

$$u = 4 \quad v = x^6$$

$$\frac{du}{dx} = 0 \quad \frac{dv}{dx} = 6x^5$$

and

$$\begin{aligned}
 \frac{dy}{dx} &= \frac{(x^6)(0) - 4(6x^5)}{(x^6)^2} \\
 &= \frac{-24x^5}{x^{12}} \\
 &= \frac{-24}{x^7}
 \end{aligned}$$

NOTE: This example could also be solved using Rule 2, as follows:

$$\begin{aligned}
 y &= 4x^{-6} \\
 \frac{dy}{dx} &= -24x^{-7} \\
 &= \frac{-24}{x^7}
 \end{aligned}$$

EXAMPLE

If $y = \frac{x^3 + 16}{x^2}$,

$$\frac{d}{dx}(x^3 + 16) = 3x^2$$

$$\frac{d}{dx}(x^2) = 2x$$

and

$$\begin{aligned}
 \frac{dy}{dx} &= \frac{(x^2)(3x^2) - (x^3 + 16)(2x)}{x^4} \\
 &= \frac{3x^4 - 2x^4 - 32x}{x^4} \\
 &= \frac{x^4 - 32x}{x^4} \\
 &= \frac{x^3 - 32}{x^3}
 \end{aligned}$$

NOTE: This example could also be solved using Rule 2, as follows:

$$y = x + 16x^{-2}$$

$$\frac{dy}{dx} = 1 - 32x^{-3}$$

$$= 1 - \frac{32}{x^3}$$

$$= \frac{x^3 - 32}{x^3}$$

EXAMPLE

$$\text{If } y = \frac{x^3 + 3x + 15}{x^2 + 2x + 2},$$

$$\frac{d}{dx}(x^3 + 3x + 15) = 3x^2 + 3$$

$$\frac{d}{dx}(x^2 + 2x + 2) = 2x + 2$$

and

$$\begin{aligned} \frac{dy}{dx} &= \frac{(x^2 + 2x + 2)(3x^2 + 3) - (x^3 + 3x + 15)(2x + 2)}{(x^2 + 2x + 2)^2} \\ &= \frac{3x^4 + 6x^3 + 6x^2 + 3x^2 + 6x + 6 - 2x^4 - 6x^2 - 30x - 2x^3 - 6x - 30}{(x^2 + 2x + 2)^2} \\ &= \frac{x^4 + 4x^3 + 3x^2 - 30x - 24}{(x^2 + 2x + 2)^2} \end{aligned}$$

Rule 7. The derivative of the n th power of a differentiable function is the product of n , the $(n - 1)$ st power of the function, and the derivative of the function: If $y = u^n$, where $u = f(x)$ is a differentiable function of x and n is any real number (positive or negative, integer or noninteger),

$$\frac{dy}{dx} = nu^{n-1} \frac{du}{dx}$$

Special Case: If $u = f(x) = x$, then $u^n = x^n$ and

$$\frac{dy}{dx} = nu^{n-1} = nx^{n-1}$$

which is Rule 2.

EXAMPLE

If $y = (x^2 + 3)^3$,

$$\frac{d}{dx}(u^n) = nu^{n-1} \frac{du}{dx}$$

$$u = (x^2 + 3)$$

$$n = 3$$

$$\frac{du}{dx} = 2x$$

and

$$\frac{dy}{dx} = 3(x^2 + 3)^2(2x)$$

$$= 6x(x^2 + 3)^2$$

EXAMPLE

If $y = (x + 3)^{-1/3}$,

$$\frac{d}{dx}(u^n) = nu^{n-1} \frac{du}{dx}$$

$$u = (x + 3)$$

$$n = -\frac{1}{3}$$

$$\frac{du}{dx} = 1$$

and

$$\frac{dy}{dx} = \left(-\frac{1}{3}\right)(x + 3)^{-4/3}(1)$$

$$= \frac{-1}{3(x + 3)^{4/3}}$$

EXAMPLE

If $y = [(x + 1)/(x - 1)]^2$,

$$\frac{dy}{dx}(u^n) = nu^{n-1} \frac{du}{dx}$$

but note that $u = (x + 1)/(x - 1)$ is a fraction and thus Rule 6 must be used to obtain dx .

$$\frac{du}{dx} = \frac{(x - 1) - (x + 1)}{(x - 1)^2}$$

$$= \frac{-2}{(x - 1)^2}$$

and

$$\begin{aligned}\frac{dy}{dx} &= 2\left(\frac{x+1}{x-1}\right)\left(\frac{-2}{(x-1)^2}\right) \\ &= \frac{-4(x+1)}{(x-1)^3}\end{aligned}$$

EXAMPLE

If $y = (x+1)^2(x^2+1)^{-3}$,

$$\frac{dy}{dx}(uv) = u \frac{dv}{dx} + v \frac{du}{dx}$$

but note that $u = (x+1)^2$ and $v = (x^2+1)^{-3}$ and thus Rule 7 must be used to obtain $\frac{du}{dx}$ and

$$\frac{dv}{dx}.$$

$$\frac{d}{dx}(x+1)^2 = 2(x+1)$$

$$\begin{aligned}\frac{d}{dx}(x^2+1)^{-3} &= -3(x^2+1)^{-4}(2x) \\ &= -6x(x^2+1)^{-4}\end{aligned}$$

and

$$\begin{aligned}\frac{dy}{dx} &= (x+1)^2(-6x)(x^2+1)^{-4} + (x^2+1)^{-3}(2)(x+1) \\ &= (x+1)(x^2+1)^{-4}[-6x(x+1) + 2(x^2+1)] \\ &= (x+1)(x^2+1)^{-4}(-4x^2 - 6x + 2) \\ &= \frac{-2(x+1)(2x^2+3x-1)}{(x^2+1)^4}\end{aligned}$$

EXAMPLE

If $y = (x^2-x)^{-2}$,

$$\begin{aligned}\frac{dy}{dx} &= -2(x^2-x)^{-3}(2x-1) \\ &= -\frac{2(2x-1)}{(x^2-x)^3} \\ &= \frac{2(1-2x)}{x^3(x-1)^3}\end{aligned}$$

EXAMPLE

If $y = x^2(x+1)^{-1}$,

$$\begin{aligned}
 \frac{dy}{dx} &= 2x(x+1)^{-1} + x^2(-1)(x+1)^{-2} \\
 &= (x+1)^{-2}[2x(x+1) - x^2] \\
 &= (x+1)^{-2}(2x^2 + 2x - x^2) \\
 &= (x+1)^{-2}(x^2 + 2x) \\
 &= \frac{x(x+2)}{(x+1)^2}
 \end{aligned}$$

EXAMPLE

If $y = (x + x^{-1})^2$,

$$\begin{aligned}
 \frac{dy}{dx} &= 2(x + x^{-1})(1 - x^{-2}) \\
 &= 2(x - x^{-1} + x^{-1} - x^{-3}) \\
 &= 2(x - x^{-3}) \\
 &= 2x^{-3}(x^4 - 1) \\
 &= \frac{2(x^4 - 1)}{x^3}
 \end{aligned}$$

EXAMPLE

If $y = 2x(3x^2 + 1)^{-1}$,

$$\begin{aligned}
 \frac{dy}{dx} &= 2(3x^2 + 1)^{-1} + 2x(-1)(3x^2 + 1)^{-2}(6x) \\
 &= 2(3x^2 + 1)^{-2}(3x^2 + 1 - 6x^2) \\
 &= 2(3x^2 + 1)^{-2}(1 - 3x^2) \\
 &= \frac{2(1 - 3x^2)}{(3x^2 + 1)^2}
 \end{aligned}$$

PROBLEMS

Find the first derivative with respect to x for each of the following functions $y = f(x)$.

1. $y = 2x^3 + 4x^2 - 5x + 8$
2. $y = -5 + 3x - \frac{3}{2}x^2 - 7x^3$
3. $y = (2x^2 + 4x - 5)^6$
4. $y = \frac{1}{3}x^{5/2} + \frac{1}{3}x^{3/2}$
5. $y = (1 - x^2)^{1/2}$
6. $y = \frac{6}{x} + \frac{4}{x^2} - \frac{3}{x^3}$
7. $y = (x^3 - 3x)^4$
8. $y = (x + x^{-1})^2$
9. $y = (x - 1)^3(x + 2)^4$
10. $y = (x + 2)^2(2 - x)^3$
11. $y = (x + 1)^2(x^2 + 1)^{-3}$
12. $y = \frac{2x + 1}{x^2 - 1}$

13. $y = \frac{x}{x^2 + 1}$
14. $y = \left(\frac{x+1}{x-1}\right)^2$
15. $y = (x^2 - x)^{-2}$
16. $y = x^2(x+1)^{-1}$
17. $y = \frac{1}{(x^2 - 9)^{1/2}}$
18. $y = \frac{1}{(16 - x^2)^{1/2}}$
19. $y = \frac{x}{(x+1)^{1/2}}$
20. $y = \frac{(x^2 + 2)^{1/2}}{x}$
21. $y = \frac{2x}{(x^5 + 10)^{1/8}}$
22. $y = (x+2)^3(x^2 + 1)^{-1}$
23. $y = \left(\frac{x^2 + 10}{x}\right)^{10}$
24. $y = \frac{3}{(x^3 - 4)^{2/3}}$
25. $y = \frac{(x^2 + 1)^{1/2}}{(2x + 4)^{1/4}}$
26. $y = (x+2)^{-3/2}(3x^2 + 1)$
27. $y = x^3(x^2 + 3)^{-1}$
28. $y = x^6 + x^{4/3} + 6x^{1/2}$

Answers to Odd-Numbered Problems

1. $6x^2 + 8x - 5$
3. $24(x+1)(2x^2 + 4x - 5)^5$
5. $-x(1 - x^2)^{-1/2}$
7. $12x^3(x^2 - 1)(x^2 - 3)^3$
9. $(x-1)^2(x+2)^3(7x+2)$
11. $-2(x+1)(x^2 + 1)^{-4}(2x^2 + 3x - 1)$
13. $\frac{1 - x^2}{(x^2 + 1)^2}$
15. $-2x^{-3}(x-1)^{-3}(2x-1)$
17. $-x(x^2 - 9)^{-3/2}$
19. $\frac{x+2}{2(x+1)^{3/2}}$
21. $(x^5 + 10)^{-9/8}(\frac{3}{4}x^5 + 20)$
23. $10\left(\frac{x^2 - 10}{x^2}\right)\left(\frac{x^2 + 10}{x}\right)^9$
25. $\frac{1}{2}(x^2 + 1)^{-1/2}(2x + 4)^{-5/4}(3x^2 + 8x - 1)$
27. $x^2(x^2 + 9)(x^2 + 3)^{-2}$

LOGARITHMIC FUNCTIONS

The notation “ln” denotes natural logarithms; the base of the natural logarithms is defined by $e = \lim_{n \rightarrow \infty} (1 + (1/n))^n$, and is approximately 2.718. The notation “log” denotes common logarithms (base 10); if any other base is meant, it is specified.

Rule 8. If $y = \log_a u$, where $u = f(x)$ is a differentiable function of x ,

$$\frac{dy}{dx} = \frac{\log_a e}{u} \frac{du}{dx}$$

EXAMPLE

$$\text{If } y = \log\left(\frac{x}{x+1}\right),$$

$$\begin{aligned} \frac{d}{dx}\left(\frac{x}{x+1}\right) &= \frac{(x+1) - x}{(x+1)^2} \\ &= \frac{1}{(x+1)^2} \end{aligned}$$

(Rule 6)

and

$$\begin{aligned}\frac{dy}{dx} &= \frac{\log e}{\frac{x}{x+1}} \left(\frac{1}{(x+1)^2} \right) \\ &= \frac{\log e}{x(x+1)}\end{aligned}$$

EXAMPLEIf $y = \sqrt{\log x}$,

$$\begin{aligned}\frac{dy}{dx} &= \frac{1}{2} (\log x)^{-1/2} \left(\frac{\log e}{x} \right) \\ &= \frac{\log e}{2x\sqrt{\log x}}\end{aligned}$$

EXAMPLEIf $y = (\log x^2)^3$,

$$\begin{aligned}\frac{dy}{dx} &= 3(\log x^2)^2 \left(\frac{\log e}{x^2} \right) (2x) \\ &= \frac{6(\log x^2)^2 \log e}{x}\end{aligned}$$

Special Case: If $y = \ln u$, where $u = f(x)$ is a differentiable function of x ,

$$\frac{dy}{dx} = \frac{1}{u} \frac{du}{dx}$$

EXAMPLEIf $y = \ln \sqrt{x^2 - 1}$,

$$\begin{aligned}\frac{d}{dx} (\sqrt{x^2 - 1}) &= \frac{1}{2} (x^2 - 1)^{-1/2} (2x) \\ &= x(x^2 - 1)^{-1/2}\end{aligned}$$

and

$$\begin{aligned}\frac{dy}{dx} &= \left(\frac{1}{\sqrt{x^2 - 1}} \right) \left(\frac{1}{2} \right) (x^2 - 1)^{-1/2} (2x) \\ &= \frac{x}{x^2 - 1}\end{aligned}$$

EXAMPLE

$$\text{If } y = \frac{\ln x}{x},$$

$$d(\ln x) = \frac{1}{x}$$

and

$$\begin{aligned} \frac{dy}{dx} &= \frac{\left(\frac{1}{x}\right)x - \ln x}{x^2} && \text{(Rule 6)} \\ &= \frac{1 - \ln x}{x^2} \end{aligned}$$

EXAMPLE

$$\text{If } y = \ln(x + \sqrt{x^2 + 4}),$$

$$\begin{aligned} \frac{d}{dx}(x + \sqrt{x^2 + 4}) &= 1 + \frac{1}{2}(x^2 + 4)^{-1/2}(2x) \\ &= 1 + x(x^2 + 4)^{-1/2} \end{aligned}$$

and

$$\begin{aligned} \frac{dy}{dx} &= \frac{1}{x + \sqrt{x^2 + 4}} [1 + x(x^2 + 4)^{-1/2}] \\ &= \frac{1 + \frac{x}{\sqrt{x^2 + 4}}}{x + \sqrt{x^2 + 4}} \\ &= \frac{x + \sqrt{x^2 + 4}}{\sqrt{x^2 + 4} (x + \sqrt{x^2 + 4})} \\ &= \frac{1}{\sqrt{x^2 + 4}} \end{aligned}$$

EXPONENTIAL FUNCTIONS

Rule 9. If $y = a^u$, where $u = f(x)$ is a differentiable function of x ,

$$\frac{dy}{dx} = a^u \ln a \frac{du}{dx}$$

EXAMPLE

$$\text{If } y = 2^{-x},$$

$$\begin{aligned} \frac{dy}{dx} &= 2^{-x}(\ln 2)(-1) \\ &= -2^{-x} \ln 2 \end{aligned}$$

EXAMPLE

If $y = 10^{x^2-x}$,

$$\begin{aligned}\frac{dy}{dx} &= 10^{x^2-x}(\ln 10)(2x-1) \\ &= (2x-1)10^{x^2-x} \ln 10\end{aligned}$$

EXAMPLE

If $y = a^x x^a$,

$$\begin{aligned}\frac{dy}{dx} &= a^x(\ln a)x^a + a^x a x^{a-1} \\ &= a^x x^{a-1}(a + x \ln a)\end{aligned}$$

Special Case: If $y = e^u$, where $u = f(x)$ is a differentiable function of x ,

$$\frac{dy}{dx} = e^u \frac{du}{dx}$$

EXAMPLE

If $y = \frac{e^x}{x}$,

$$\begin{aligned}\frac{dy}{dx} &= \frac{xe^x - e^x}{x^2} \\ &= \frac{e^x(x-1)}{x^2}\end{aligned}$$

EXAMPLE

If $y = 10e^{x^2+4}$,

$$\begin{aligned}\frac{dy}{dx} &= 10e^{x^2+4}(2x) \\ &= 20xe^{x^2+4}\end{aligned}$$

EXAMPLE

If $y = \frac{e^x - e^{-x}}{e^x + e^{-x}}$,

$$\begin{aligned}\frac{d}{dx}(e^x - e^{-x}) &= e^x + e^{-x} \\ \frac{d}{dx}(e^x + e^{-x}) &= e^x - e^{-x}\end{aligned}$$

and

$$\begin{aligned}\frac{dy}{dx} &= \frac{(e^x + e^{-x})^2 - (e^x - e^{-x})^2}{(e^x + e^{-x})^2} \\ &= \frac{e^{2x} + 2 + e^{-2x} - e^{2x} + 2 - e^{-2x}}{(e^x + e^{-x})^2} = \frac{4}{(e^x + e^{-x})^2}\end{aligned}$$

Rule 10. If $y = u^v$, where $u = f(x)$ and $v = g(x)$ are differentiable functions of x , differentiate $\ln y = v \ln u$ to obtain

$$\frac{dy}{dx} = vu^{v-1} \frac{du}{dx} + u^v \ln u \frac{dv}{dx}$$

Note that Rule 9 is a special case of Rule 10, where $u = a$. Some algebraic functions are also easier to differentiate if the logarithm of each side is differentiated and the resulting equation is solved for the derivative of the function. (See the fourth example below.)

EXAMPLE

If $y = x^{x^2}$,

$$\ln y = x^2 \ln x$$

$$\frac{1}{y} \frac{dy}{dx} = 2x \ln x + x^2 \left(\frac{1}{x} \right)$$

$$= x(2 \ln x + 1)$$

$$\frac{dy}{dx} = x^{x^2+1}(1 + 2 \ln x)$$

EXAMPLE

If $y = x^{e^x}$,

$$\ln y = e^x \ln x$$

$$\frac{1}{y} \frac{dy}{dx} = e^x \ln x + e^x \left(\frac{1}{x} \right)$$

$$= e^x \left(\ln x + \frac{1}{x} \right)$$

$$\frac{dy}{dx} = x^{e^x} e^x \left(\frac{1}{x} + \ln x \right)$$

EXAMPLE

If $y = (x^4 + 5)^{x^2}$,

$$\ln y = x^2 \ln (x^4 + 5)$$

$$\frac{1}{y} \frac{dy}{dx} = 2x \ln (x^4 + 5) + x^2 \left(\frac{4x^3}{x^4 + 5} \right)$$

$$= 2x \left[\ln (x^4 + 5) + \frac{2x^4}{x^4 + 5} \right]$$

$$\frac{dy}{dx} = 2x(x^4 + 5)^{x^2-1} [(x^4 + 5) \ln (x^4 + 5) + 2x^4]$$

EXAMPLE

If $y = \sqrt[3]{\frac{x+1}{x-1}}$,

$$\ln y = \frac{1}{3} \ln (x+1) - \frac{1}{3} \ln (x-1)$$

$$\frac{1}{y} \frac{dy}{dx} = \frac{1}{3(x+1)} - \frac{1}{3(x-1)}$$

$$= \frac{x-1-(x+1)}{3(x^2-1)}$$

$$= \frac{-2}{3(x^2-1)}$$

$$\frac{dy}{dx} = \frac{-2}{3(x^2-1)} \left(\frac{x+1}{x-1} \right)^{1/3}$$

PROBLEMS

Find the first derivative of each of the following functions.

1. $y = \log(1-2t)$

13. $y = e^{-1/x}$

2. $y = \ln \left(\frac{1-4x}{1+4x} \right)$

14. $y = x^{x^3}$

15. $y = x^{e^{x^2+1}}$

3. $R = \log_a (a^2 - x^2)^3$

16. $y = \frac{(x+1)^6}{(x^2+2x+2)^3}$

4. $t = 6^{-2u}$

17. $y = 2x^2 e^{x^2+3}$

5. $y = e^x \ln x$

18. $y = \log_5 (x^3 + x^2)^6$

6. $y = \frac{\sqrt{x+1}}{\sqrt[3]{x+2}}$

19. $y = x^{2x^4+x}$

7. $t = e^{\ln x}$

20. $y = \log_b (b - x^3)^2$

8. $y = a^x e^x$

21. $y = \left(\frac{x+3}{x-1} \right)^2$

9. $y = \log (x^3 - 3x)^{1/3}$

22. $y = e^{\ln x^2}$

10. $y = t^2 \ln t$

23. $y = 16^{x^2-2x}$

11. $y = 2 \ln \left(\sqrt{\frac{1-t^2}{t}} \right)$

24. $y = x^{x^4}$

12. $y = 25^{3x^3-6}$

25. $y = e^{x^2+4x+3}$
 26. $y = e^{\ln(x+3)}$
 27. $y = \log_8 (6 - x^2)^4$
 28. $y = x^{x^2+2}$
 29. $y = 2 \ln (x^3 + 4x^2)^{1/4}$
 30. $y = (x^2 + 4)^2 e^{x^2+1}$
 31. $y = \frac{1}{3} e^{3 \ln x}$
 32. $y = x e^{\ln (x^2+5)}$
 33. $y = x^2 e^{x^2+4x+2}$
 34. $y = e^{\ln (x^2+3x^2+10)}$
 35. $y = (x+1)^x$
 36. $y = (x+1)^{x^2+1}$

Answers to Odd-Numbered Problems

1. $\frac{-2 \log e}{1-2t}$
 3. $\frac{6x \log_a e}{x^2 - a^2}$
 5. $\frac{e^x}{x} (1 + x \ln x)$
 7. 1
 9. $\frac{(x^2 - 1) \log e}{x(x^2 - 3)}$
 11. $\frac{t^2 + 1}{t(t^2 - 1)}$
 13. $\frac{1}{x^2} e^{-1/x}$
 15. $x^{e^{x^2+1}} (e^{x^2+1}) \left(\frac{1}{x} + 2x \ln x \right)$
 17. $4x(x^2 + 1)e^{x^2+3}$
 19. $x^{2x^2+x} [2x^3 + 1 + (8x^3 + 1) \ln x]$
 21. $\frac{-8(x+3)}{(x-1)^3}$
 23. $2(x-1)16^{x^2-2x} \ln 16$
 25. $(2x+4)e^{x^2+4x+3}$
 27. $\frac{-8x \log_8 e}{6 - x^2}$
 29. $\frac{3x+8}{2(x^2+4x)}$
 31. x^2
 33. $2x(x+1)^2 e^{x^2+4x+2}$
 35. $(x+1)^{x-1} [x + (x+1) \ln (x+1)]$

TRIGONOMETRIC FUNCTIONS

Rule 11. If $y = \sin u$, where $u = f(x)$ is a differentiable function of x ,

$$\frac{dy}{dx} = \cos u \frac{du}{dx}$$

EXAMPLE

If $y = \sin (x^2 + \theta)$,

$$\frac{dy}{dx} = 2x \cos (x^2 + \theta)$$

EXAMPLE

If $y = \sin^2 3x$,

$$\begin{aligned} \frac{dy}{dx} &= 2(\sin 3x)(\cos 3x)(3) \\ &= 6 \sin 3x \cos 3x \end{aligned}$$

Rule 12. If $y = \cos u$, where $u = f(x)$ is a differentiable function of x ,

$$\frac{dy}{dx} = -\sin u \frac{du}{dx}$$

Rule 13. If $y = \tan u$, where $u = f(x)$ is a differentiable function of x ,

$$\frac{dy}{dx} = \sec^2 u \frac{du}{dx}$$

Similarly,

$$y = \csc u = \frac{1}{\sin u}$$

$$y = \sec u = \frac{1}{\cos u}$$

$$y = \cot u = \frac{1}{\tan u}$$

can be differentiated to obtain the derivatives of the cosecant, secant, and cotangent functions.

$$\frac{d}{dx}(\csc u) = -\csc u \cot u \frac{du}{dx}$$

$$\frac{d}{dx}(\sec u) = \sec u \tan u \frac{du}{dx}$$

$$\frac{d}{dx}(\cot u) = -\csc^2 u \frac{du}{dx}$$

EXAMPLE

If $y = \cos x + \sec x$,

$$\begin{aligned} \frac{dy}{dx} &= -\sin x + \sec x \tan x \\ &= \sin x(\sec^2 x - 1) \end{aligned}$$

EXAMPLE

If $y = (1 + \cot x)^2$,

$$\begin{aligned} \frac{dy}{dx} &= 2(1 + \cot x)(-\csc^2 x) \\ &= -2 \csc^2 x(1 + \cot x) \end{aligned}$$

EXAMPLE

$$\text{If } y = \tan^2 (x + 3) + \sec^2 (x + 3),$$

$$\begin{aligned}\frac{dy}{dx} &= 2 \tan (x + 3) \sec^2 (x + 3) + 2 \sec^2 (x + 3) \tan (x + 3) \\ &= 4 \tan (x + 3) \sec^2 (x + 3)\end{aligned}$$

EXAMPLE

$$\text{If } y = (\sin x + \cos x)^2,$$

$$\begin{aligned}\frac{dy}{dx} &= 2(\sin x + \cos x)(\cos x - \sin x) \\ &= 2(\cos^2 x - \sin^2 x)\end{aligned}$$

EXAMPLE

$$\text{If } y = \frac{\cot (x^2 + 2)}{\sec (x^2 + 2)},$$

$$\begin{aligned}\frac{dy}{dx} &= \frac{-2x \csc^2 (x^2 + 2) \sec (x^2 + 2) - 2x \cot (x^2 + 2) \sec (x^2 + 2) \tan (x^2 + 2)}{\sec^2 (x^2 + 2)} \\ &= \frac{-2x[\csc^2 (x^2 + 2) + 1]}{\sec^2 (x^2 + 2)}\end{aligned}$$

EXAMPLE

$$\text{If } y = \tan (x^3 + 2x + 3) \sin (x^3 + 2x + 3),$$

$$\begin{aligned}\frac{dy}{dx} &= (3x^2 + 2) \sec^2 (x^3 + 2x + 3) \sin (x^3 + 2x + 3) \\ &\quad + (3x^2 + 2) \tan (x^3 + 2x + 3) \cos (x^3 + 2x + 3) \\ &= (3x^2 + 2) \sin (x^3 + 2x + 3)[\sec^2 (x^3 + 2x + 3) + 1]\end{aligned}$$

PROBLEMS

Find the first derivative of each of the following functions.

$$1. y = \frac{\csc x}{x}$$

$$5. y = \sin 3x \tan 3x$$

$$2. y = \frac{\sin x - \cos x}{x}$$

$$6. y = \cot (x^2 + 1)$$

$$3. y = x^2 \cot x$$

$$7. y = x + \cot x$$

$$8. y = \frac{x + \sin x}{x}$$

$$4. y = \tan x \csc x$$

9. $y = \cos^2 2x$
10. $y = \sin x \cos x$
11. $y = \sin^2 x + \cos^2 x$
12. $y = x \sin x$
13. $y = \frac{\sin x}{x}$
14. $y = x^2 - \cos x$
15. $y = \frac{\tan x}{(1+x)^2}$
16. $y = \sin x^2$
17. $y = \tan 3x$
18. $y = \tan x + \sec 2x$
19. $y = \frac{\sin x}{\sec x}$
20. $y = \cos x \csc x$
21. $y = \sec^2 3x$
22. $y = \frac{x}{\sec x}$
23. $y = \frac{1}{1 + \cot x}$
24. $y = \frac{\cot x}{1 + x^2}$

Answers to Odd-Numbered Problems

1. $\frac{-\csc x(1+x \cot x)}{x^2}$
3. $2x \cot x - x^2 \csc^2 x$
5. $3 \sin 3x(1 + \sec^2 3x)$
7. $1 - \csc^2 x$
9. $-4 \cos 2x \sin 2x$
11. 0
13. $\frac{x \cos x - \sin x}{x^2}$
15. $\frac{(1+x) \sec^2 x - 2 \tan x}{(1+x)^3}$
17. $3 \sec^2 3x$
19. $\cos^2 x - \sin^2 x$
21. $6 \sec^2 3x \tan 3x$
23. $\frac{1}{(\sin x + \cos x)^2}$

INVERSE FUNCTIONS

Rule 14. The derivative of the inverse of a function is equal to the reciprocal of the derivative of the function. If $y = f(x)$ and $x = g(y)$ are inverse differentiable functions,

$$\frac{dy}{dx} = \frac{1}{\frac{dx}{dy}}$$

or

$$\frac{df(x)}{dx} = \frac{1}{\frac{dg(y)}{dy}}$$

EXAMPLE

If $x = y + \frac{1}{3}y^3 + \frac{1}{5}y^5$, find $\frac{dy}{dx}$.

$$\frac{dx}{dy} = 1 + y^2 + y^4$$

$$\frac{dy}{dx} = \frac{1}{\frac{dx}{dy}} = \frac{1}{1 + y^2 + y^4}$$

EXAMPLE

If $x = \ln(y^4 + 3y^2)$,

$$\begin{aligned}\frac{dx}{dy} &= \frac{4y^3 + 6y}{y^4 + 3y^2} \\ &= \frac{4y^2 + 6}{y^3 + 3y^2} \\ \frac{dy}{dx} &= \frac{1}{\frac{dx}{dy}} = \frac{y^3 + 3y^2}{4y^2 + 6}\end{aligned}$$

EXAMPLE

If $x = e^{-y^2 + 4y + 2}$,

$$\begin{aligned}\frac{dx}{dy} &= (-2y + 4)e^{-y^2 + 4y + 2} \\ \frac{dy}{dx} &= \frac{1}{\frac{dx}{dy}} = \frac{e^{y^2 - 4y - 2}}{4 - 2y}\end{aligned}$$

EXAMPLE

If $x = \tan^2(y^3 + a)$,

$$\begin{aligned}\frac{dx}{dy} &= 6y^2 \tan(y^3 + a) \sec^2(y^3 + a) \\ \frac{dy}{dx} &= \frac{1}{\frac{dx}{dy}} = \frac{1}{6y^2} \cot(y^3 + a) \cos^2(y^3 + a)\end{aligned}$$

PROBLEMS

Find the first derivative with respect to x of each of the following functions $x = f(y)$.

- | | |
|------------------------------|----------------------------|
| 1. $x = y^2 + 3y + 2$ | 4. $x = \ln(y^5 - 6)$ |
| 2. $x = y^3 - 2y^{3/2} + 6y$ | 5. $x = \cos^2 y + \sin y$ |
| 3. $x = 5y^4 - e^y$ | 6. $x = y \sec y$ |

Find the first derivative with respect to y of each of the following functions $y = f(x)$.

- | | |
|--|----------------------------|
| 7. $y = e^{-x} - 6$ | 9. $y = \tan x \sin x$ |
| 8. $y = \ln\left(\frac{x^3}{x+2}\right)$ | 10. $y = \frac{x}{\cos x}$ |

Answers to Odd-Numbered Problems

1. $\frac{1}{2y+3}$ 5. $\frac{\sec y}{1-2\sin y}$
 3. $\frac{1}{20y^3 - e^y}$ 7. $-e^x$
 9. $\frac{\csc x}{1 + \sec^2 x}$

COMPOSITE FUNCTIONS

If y is a function of u and u is a function of x , then y is a function of a function or a *composite function*; the derivative of y with respect to x is the product of the derivative of y with respect to u and the derivative of u with respect to x .

Rule 15. If $y = \Phi(u)$ and $u = F(x)$, that is, $y = \Phi[F(x)] = f(x)$,

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$$

(This is sometimes called the *chain rule of differentiation*.)

Note that Rules 7, 8, 9, 10, 11, 12, and 13 are special cases of Rule 15.

EXAMPLE

If $y = u^{2/3}$ and $u = x^2 + 1$, find $\frac{dy}{dx}$.

$$\frac{dy}{du} = \frac{2}{3}u^{-1/3}$$

$$= \frac{2}{3(x^2 + 1)^{1/3}}$$

$$\frac{du}{dx} = 2x$$

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx} = \frac{4x}{3(x^2 + 1)^{1/3}}$$

Alternatively,

$$y = u^{2/3} = (x^2 + 1)^{2/3}$$

$$\frac{dy}{dx} = \frac{2}{3}(x^2 + 1)^{-1/3}(2x)$$

$$= \frac{4x}{3(x^2 + 1)^{1/3}}$$

EXAMPLE

If $y = \frac{z^2}{z^2 + 1}$ and $z = \sqrt{2u + 1}$, find $\frac{dy}{du}$.

$$\frac{dy}{dz} = \frac{2z(z^2 + 1) - z^2(2z)}{(z^2 + 1)^2}$$

$$= \frac{2z}{(z^2 + 1)^2}$$

$$= \frac{2\sqrt{2u+1}}{(2u+2)^2}$$

$$= \frac{\sqrt{2u+1}}{2(u+1)^2}$$

$$\frac{dz}{du} = \frac{1}{2}(2u+1)^{-1/2}(2)$$

$$= \frac{1}{\sqrt{2u+1}}$$

$$\frac{dy}{du} = \frac{dy}{dz} \cdot \frac{dz}{du} = \frac{1}{2(u+1)^2}$$

Alternatively,

$$y = \frac{z^2}{z^2 + 1} = \frac{2u + 1}{2u + 2}$$

$$\frac{dy}{dx} = \frac{2(2u+2) - (2u+1)(2)}{(2u+2)^2}$$

$$= \frac{2}{(2u+2)^2}$$

$$= \frac{1}{2(u+1)^2}$$

EXAMPLE

If $y = \ln(w^3 - 1)$ and $w = z^2 - 1$, find $\frac{dy}{dz}$.

$$\frac{dy}{dw} = \frac{3w^2}{w^3 - 1} = \frac{3(z^2 - 1)^2}{(z^2 - 1)^3 - 1}$$

$$\frac{dw}{dz} = 2z$$

$$\frac{dy}{dz} = \frac{dy}{dw} \cdot \frac{dw}{dz} = \frac{6z(z^2 - 1)^2}{(z^2 - 1)^3 - 1}$$

Alternatively,

$$y = \ln(w^3 - 1) = \ln[(z^2 - 1)^3 - 1]$$

$$\frac{dy}{dz} = \frac{3(z^2 - 1)^2(2z)}{(z^2 - 1)^3 - 1}$$

$$= \frac{6z(z^2 - 1)^2}{(z^2 - 1)^3 - 1}$$

PROBLEMS

Determine $\frac{dy}{dx}$ for each of the following functions.

1. $x = \sqrt{1-2y}$
2. $x = \sqrt[3]{4-9y}$
3. $x = (2-3y^2)^3$
4. $x = \left(a - \frac{b}{y}\right)^2$
5. $x = y\sqrt{a^2 + y^2}$
6. $x = \frac{a-y}{a+y}$
7. $x = y\sqrt{a+by}$
8. $x = \frac{b}{a}\sqrt{a^2 - y^2}$
9. $x = \ln(ay^2 + b)$
10. $x = \frac{e^y - e^{-y}}{e^y + e^{-y}}$
11. $x = \frac{2}{e^y}$
12. $x = \frac{\ln y}{y}$
13. $x = y^2 e^{-y}$
14. $x = \ln(y^2 e^y)$
15. $x = \ln(\sqrt{\cos 2y})$
16. $x = \frac{4}{\sqrt{\sec y}}$
17. $x = \ln\left(\frac{1 + \sin y}{1 - \sin y}\right)^{1/2}$
18. $x = (\sin 2y)^{1/2}$
19. $y = u^6; u = 1 + 2\sqrt{x}$
20. $y = u \sin u; u = \ln x$
21. $y = u^2 \cos u; u = ax^2$
22. $y = \frac{a-u}{a+u}; u = \frac{b-x}{b+x}$
23. If $y = x^4 + 5$ and $x = \log z$, find $\frac{dy}{dz}$.
24. If $y = e^{3u}$ and $u = 2x^2 - 3x$, find $\frac{dy}{dx}$.
25. If $u = \ln(y+4)$ and $y = x^2$, find $\frac{du}{dx}$.
26. If $x = \frac{4y+2}{y+6}$ and $y = u^3 + 10u$, find $\frac{dx}{du}$.
27. If $y = e^t + 6$ and $t = \ln(x^2 + 6x)$, find $\frac{dy}{dx}$.
28. If $y = \frac{4t-8}{t+4}$ and $t = x^2 - 4$, find $\frac{dy}{dx}$.

Answers to Odd-Numbered Problems

1. $-x$
3. $-\frac{1}{18y(2-3y^2)^2}$
5. $\frac{(a^2 + y^2)^{1/2}}{a^2 + 2y^2}$
7. $\frac{2(a+by)^{1/2}}{2a+3by}$
9. $\frac{ay^2 + b}{2ay}$
11. $-\frac{1}{2}e^y$
13. $\frac{e^y}{y(2-y)}$
15. $-\cot 2y$
17. $\cos y$
19. $\frac{6(1+2\sqrt{x})^5}{\sqrt{x}}$
21. $2a^2x^3(2\cos ax^2 - ax^2 \sin ax^2)$
23. $\frac{4(\log z)^3}{z} \log e$
25. $\frac{2x}{x^2 + 4}$
27. $2x + 6$

Summary of Procedure for Differentiation

1. Determine whether the entire function to be differentiated is a sum or difference of functions; a product or quotient of functions; a logarithmic, exponential, or trigonometric function of a function; a power of a function; a composite function; or some combination of these.
2. Differentiate using the appropriate rule for the entire function and the appropriate rules for the different parts of the function.

EXAMPLE

Find the first derivative of $y = x^2 + \ln x^3$.

The entire function is the sum of two functions

$$y = u(x) + v(x) \quad \text{where } u(x) = x^2 \\ v(x) = \ln x^3$$

$$\text{Thus } \frac{dy}{dx} = \frac{du}{dx} + \frac{dv}{dx} \quad (\text{Rule 4}).$$

$u = x^2$ is differentiated using Rule 2, and $v = \ln x^3$ is differentiated using Rules 2 and 8.

$$\frac{du}{dx} = 2x \quad \text{and} \quad \frac{dv}{dx} = \frac{3}{x}$$

Thus

$$\frac{dy}{dx} = 2x + \frac{3}{x}$$

EXAMPLE

Find the first derivative of $y = e^{x^2+2}(x+1)^2$.

The entire function is the product of two functions

$$y = u(x) \cdot v(x) \quad \text{where } u(x) = e^{x^2+2} \\ v(x) = (x+1)^2$$

$$\text{Thus } \frac{dy}{dx} = u \frac{dv}{dx} + v \frac{du}{dx} \quad (\text{Rule 5}).$$

$u = e^{x^2+2}$ is differentiated using Rules 1, 2, 4, and 9, and $v = (x+1)^2$ is differentiated using Rules 1, 2, 4, and 7.

$$\frac{du}{dx} = 2xe^{x^2+2}$$

$$\frac{dv}{dx} = 2(x+1)$$

Thus

$$\begin{aligned} \frac{dy}{dx} &= 2e^{x^2+2}(x+1) + 2x(x+1)^2e^{x^2+2} \\ &= 2(x+1)e^{x^2+2}(1+x^2+x) \\ &= 2(x+1)(x^2+x+1)e^{x^2+2} \end{aligned}$$

EXAMPLE

Find the first derivative of $y = x^{\tan x}$.

The entire function is a function raised to a functional power

$$y = u(x)^{v(x)} \quad \text{where } u(x) = x$$

$$v(x) = \tan x$$

$$\text{Thus } \frac{dy}{dx} = vu^{v-1} \frac{du}{dx} + (\ln u)u^v \frac{dv}{dx} \quad (\text{Rule 10}).$$

$u = x$ is differentiated using Rule 2 and $v = \tan x$ is differentiated using Rule 13.

$$\frac{du}{dx} = 1 \quad \text{and} \quad \frac{dv}{dx} = \sec^2 x$$

Thus

$$\begin{aligned} \frac{dy}{dx} &= (\tan x)x^{(\tan x)-1} + (\ln x)x^{\tan x}(\sec^2 x) \\ &= x^{\tan x} \left(\frac{\tan x}{x} + \sec^2 x \ln x \right) \end{aligned}$$

PROBLEMS

Obtain the first derivative of each of the following functions.

1. $y = \frac{2}{x+1}$

2. $s = \frac{\theta+4}{\theta}$

3. $y = \ln(x + \sqrt{1+x^2})$

4. $y = \frac{2}{e^x}$

5. $y = \frac{x^3+1}{x}$

6. $y = \frac{3}{x^2+2}$

7. $s = \frac{1}{1-2t}$

8. $y = \frac{x^2}{4-x^2}$

9. $y = e^{ax} \sin bx$

10. $y = \frac{x}{x^2+1}$

11. $y = \frac{ax+b}{bx+d}$

12. $y = \ln\left(\tan \frac{x}{2}\right)$

13. $y = \ln^3 x$

14. $y = \log\left(\frac{2}{x}\right)$

15. $y = 10^{nx}$

16. $y = \ln(ax^2 + b)$

17. $s = e^{-t} \cos 2t$

18. $y = \frac{1}{3} \tan^3 \theta - \tan \theta + \theta$

19. $s = \frac{a+bt+ct^2}{\sqrt{t}}$

20. $z = \frac{a+bx+cx^2}{x}$

21. $z = a^{2y}$

22. $y = e^{x^2}$

23. $y = x^x$

24. $y = x^n(a+bx)^m$

25. $y = x^{\sin x}$

26. $y = x^{\sqrt{x}}$

27. $s = \left(\frac{a}{t}\right)^t$

28. $y = (\cos x)^x$

Answers to Odd-Numbered Problems

1. $-\frac{2}{(x+1)^2}$
3. $\frac{1}{\sqrt{1+x^2}}$
5. $\frac{2x^3-1}{x^2}$
7. $\frac{2}{(1-2t)^2}$
9. $e^{ax}(a \sin bx + b \cos bx)$
11. $\frac{ad-bc}{(cx+d)^2}$
13. $\frac{3 \ln^2 x}{x}$
15. $10^{nx} n \ln 10$
17. $-e^{-t}(\cos 2t + 2 \sin 2t)$
19. $-\frac{1}{2}at^{-3/2} + \frac{1}{2}bt^{-1/2} + \frac{3}{2}ct^{1/2}$
21. $2a^{2y} \ln a$
23. $x^x(1 + \ln x)$
25. $x^{\sin x} \left(\frac{\sin x}{x} + \cos x \ln x \right)$
27. $\left(\frac{a}{t} \right)' \left(\ln \left(\frac{a}{t} \right) - 1 \right)$

SUMMARY OF RULES FOR DIFFERENTIATION

Algebraic Functions

Rule 1. If $y = c$, $\frac{dy}{dx} = 0$.

Rule 2. If $y = x^n$, $\frac{dy}{dx} = nx^{n-1}$.

Rule 3. If $y = cu$, where $u = f(x)$, $\frac{dy}{dx} = c \frac{du}{dx}$.

Rule 4. If $y = u + v$, where $u = f(x)$ and $v = g(x)$,

$$\frac{dy}{dx} = \frac{du}{dx} + \frac{dv}{dx}$$

Rule 5. If $y = uv$, where $u = f(x)$ and $v = g(x)$,

$$\frac{dy}{dx} = u \frac{dv}{dx} + v \frac{du}{dx}$$

Rule 6. If $y = \frac{u}{v}$, where $u = f(x)$ and $v = g(x)$,

$$\frac{dy}{dx} = \frac{v \frac{du}{dx} - u \frac{dv}{dx}}{v^2}$$

Rule 7. If $y = u^n$, where $u = f(x)$,

$$\frac{dy}{dx} = nu^{n-1} \frac{du}{dx}$$

Special Case: If $u = f(x) = x$, then $u^n = x^n$ and $\frac{dy}{dx} = nu^{n-1} = nx^{n-1}$ (Rule 3).

Logarithmic Functions

Rule 8. If $y = \log_a u$, where $u = f(x)$,

$$\frac{dy}{dx} = \frac{\log_a e}{u} \frac{du}{dx}$$

Special Case: If $a = e$, then $y = \ln u$ and $\frac{dy}{dx} = \frac{1}{u} \frac{du}{dx}$.

Exponential Functions

Rule 9. If $y = a^u$, where $u = f(x)$,

$$\frac{dy}{dx} = a^u \ln a \frac{du}{dx}$$

Special Case: If $a = e$, then $y = e^u$ and $\frac{dy}{dx} = e^u \frac{du}{dx}$.

Rule 10. If $y = u^v$, where $u = f(x)$ and $v = g(x)$,

$$\frac{dy}{dx} = vu^{v-1} \frac{du}{dx} + u^v \ln u \frac{dv}{dx}$$

Note that Rule 9 is a special case of Rule 10.

Trigonometric Functions

Rule 11. If $y = \sin u$, where $u = f(x)$,

$$\frac{dy}{dx} = \cos u \frac{du}{dx}$$

Rule 12. If $y = \cos u$, where $u = f(x)$,

$$\frac{dy}{dx} = -\sin u \frac{du}{dx}$$

Rule 13. If $y = \tan u$, where $u = f(x)$,

$$\frac{dy}{dx} = \sec^2 u \frac{du}{dx}$$

Inverse Functions

Rule 14. If $y = f(x)$ and $x = g(y)$ are inverse functions,

$$\frac{dy}{dx} = \frac{1}{\frac{dx}{dy}}$$

Composite Functions

Rule 15. If $y = \Phi(u)$ and $u = F(x)$,

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$$

Note that Rules 7, 8, 9, 10, 11, 12, and 13 are special cases of Rule 15.

2.7 Differentials

In the preceding discussion of derivatives, $\frac{dy}{dx}$ is considered not as a fraction with numerator dy and denominator dx , but as a symbol denoting the limit of $\frac{\Delta y}{\Delta x}$ as Δx approaches zero as a limit. In some problems it is useful to interpret dy and dx separately. In this context, dy is referred to as the *differential of y* and dx is referred to as the *differential of x*. Differentials are useful, for example, in applications of integral calculus and in approximating changes in the dependent variable associated with small changes in the independent variable. Approximation of the value of a function using differentials is discussed in this section.

If $f'(x)$ is the derivative of $y = f(x)$ for a particular value of x and Δx is an arbitrarily chosen increment of x , then the differential of y , denoted by $df(x)$ or dy , is

$$dy = df(x) = f'(x) \Delta x = \frac{dy}{dx} \Delta x$$

and the differential of x , denoted by dx , is

$$dx = \Delta x$$

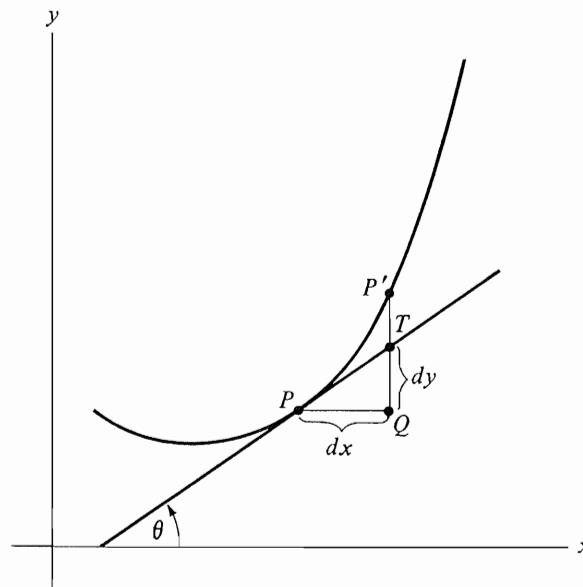
Thus the differential of a function is defined as its derivative multiplied by an increment of the independent variable. Note that if the function $f(x)$ is taken to be x , then $f'(x) = 1$ and it follows that the differential dx of x is Δx .

Geometrically, consider the curve $y = f(x)$ (see Figure 2.20) and let $f'(x)$ be the value of the derivative at P . Then $\Delta x = PQ$ and

$$dy = f'(x) \Delta x = (\tan \theta)(PQ) = \frac{QT}{PQ} PQ = QT$$

Therefore, dy is the increment (equal to QT) of the ordinate of the tangent corresponding to dx .

Figure 2.20



This geometrical argument provides the following interpretation of the derivative as a fraction. If an arbitrarily chosen increment of the independent variable x for a point $P(x, y)$ on the curve $y = f(x)$ is denoted by dx , then in the derivative

$$\frac{dy}{dx} = f'(x) = \tan \theta$$

dy denotes the corresponding increment of the ordinate of the tangent at P .

Note that the differential dy and the increment Δy of the function corresponding to the same value of $dx = \Delta x$ are not in general equal. In Figure 2.20, $dy = QT$ but $\Delta y = QP'$.

Approximations

It can be seen in Figure 2.20 that $\Delta y = QP'$ and $dy = QT$ are approximately equal when $\Delta x = PQ$ is small. In fact, the differential of a function can be made arbitrarily close to the increment of the function for sufficiently small changes in the independent variable. The value of the corresponding differential thus can be used as an approximation for the change in a function when the independent variable changes by a small amount.

EXAMPLE

If $y = x^4 - 2x^3 + 9x + 7$, use differentials to find an approximate value of y when $x = 1.997$.

Consider 1.997 to be the result of applying an increment $\Delta x = -0.003$ to an original value of $x = 2$. Then

$$\begin{aligned} dy &= (4x^3 - 6x^2 + 9) \Delta x \\ &= (32 - 24 + 9)(-0.003) \\ &= -0.051 \end{aligned}$$

Since $y = 25$ when $x = 2$ and -0.051 is the approximate change in y when x changes from 2 to 1.997,

$$\begin{aligned} y + dy &= 25 - 0.051 \\ &= 24.949 \end{aligned}$$

and y is approximately 24.949 when $x = 1.997$.

To obtain an estimate of the accuracy of this approximation, compute the error of the approximation, which is given by $\Delta y - dy$.

$$\begin{aligned} \Delta y &= f(x + \Delta x) - f(x) \\ &= (x + \Delta x)^4 - 2(x + \Delta x)^3 + 9(x + \Delta x) + 7 - x^4 + 2x^3 - 9x - 7 \\ &= x^4 - 2x^3 + 9x + 7 + 4x^3(\Delta x) + 6x^2(\Delta x)^2 \\ &\quad + 4x(\Delta x)^3 + (\Delta x)^4 - 6x^2(\Delta x) - 6x(\Delta x)^2 \\ &\quad - 2(\Delta x)^3 + 9(\Delta x) - x^4 + 2x^3 - 9x - 7 \\ &= 4x^3(\Delta x) + 6x^2(\Delta x)^2 + 4x(\Delta x)^3 + (\Delta x)^4 - 6x^2(\Delta x) \\ &\quad - 6x(\Delta x)^2 - 2(\Delta x)^3 + 9(\Delta x) \end{aligned}$$

and

$$\begin{aligned} dy &= (4x^3 - 6x^2 + 9)(\Delta x) \\ &= 4x^3(\Delta x) - 6x^2(\Delta x) + 9(\Delta x) \end{aligned}$$

Thus

$$dy - \Delta y = 6x^2(\Delta x)^2 + 4x(\Delta x)^3 + (\Delta x)^4 - 6x(\Delta x)^2 - 2(\Delta x)^3$$

And, for $x = 2$,

$$\begin{aligned} dy - \Delta y &= (24 - 12)(-0.003)^2 + (8 - 2)(-0.003)^3 + (-0.003)^4 \\ &= 0.000107838081 \end{aligned}$$

Thus $y + \Delta y = y + dy + (\Delta y - dy) = 24.949107838081$ and the approximation $y = 24.949$ is accurate to the three decimal places computed.

EXAMPLE

Using differentials, find an approximate value of $\sqrt{98}$.

The number 98 is close to the number 100, which is a perfect square. Thus the change in $y = \sqrt{x}$ corresponding to a change in x from 100 to 98 can be added to $y = \sqrt{100} = 10$ to obtain an approximate value for $\sqrt{98}$. When x changes to $x + dx$, y changes to $y + dy$, approximately.

For $x = 100$ and $\Delta x = -2$,

$$\begin{aligned} dy &= \frac{dx}{2\sqrt{x}} \\ &= \frac{-2}{2\sqrt{100}} \\ &= -0.1 \end{aligned}$$

and thus $\sqrt{98} = \sqrt{100} - 0.1 = 9.9$ approximately.

EXAMPLE

Using differentials, find an approximate value for $\tan 46^\circ$ given $\tan 45^\circ = 1$, $\sec 45^\circ = \sqrt{2}$, and $1^\circ = 0.0175$ radians.

If $y = \tan x$, $dy = \sec^2 x \, dx$. When x changes to dx , y changes to $y + dy$, approximately.

$$x = 45^\circ = \frac{\pi}{4} \text{ radians}$$

$$\Delta x = 1^\circ = 0.0175 \text{ radians}$$

$$\begin{aligned} dy &= \sec^2 x \, dx \\ &= (\sqrt{2})^2(0.0175) \\ &= 0.0350 \end{aligned}$$

Since $y = \tan 45^\circ = 1$, $y + dy = 1.0350$, and $\tan 46^\circ = 1.0350$ approximately.

Differentials can also be used to estimate small errors in calculation of a function which arise from lack of precision in the measurement of the independent variable or from other causes.

EXAMPLE

Suppose $C = 5 + 0.6x + 0.2\sqrt{x}$, where C is total consumption (billions of dollars) and x is total disposable income (billions of dollars). If $x = 25$ with a maximum error of 0.3, find the approximate maximum error in consumption.

$$\begin{aligned} dC &= 0.6 + \frac{0.1}{\sqrt{x}} dx \\ &= \left(0.6 + \frac{0.1}{\sqrt{25}}\right)(0.3) \\ &= \left(0.6 + \frac{0.1}{5}\right)(0.3) \\ &= (0.62)(0.3) \\ &= 0.186 \end{aligned}$$

If du is the error in u , then

$$\frac{du}{u} \text{ is the relative error in } u$$

and

$$100 \frac{du}{u} \text{ is the percentage error in } u$$

For this example, the approximate maximum relative error is

$$\frac{dC}{C} = \frac{0.186}{25} = 0.00744$$

and the percentage error is 0.744.

It is frequently convenient to find the relative error by logarithmic differentiation, since if $y = \ln x$, $dy = dx/x$.

EXAMPLE

For the demand function

$$x = \frac{16}{y^4}$$

where x is number of units demanded and y is price in dollars, assume y is 200 with a maximum error of 10 and determine the approximate maximum relative error in x .

$$\ln x = \ln 16 + 4 \ln y$$

$$\frac{dx}{x} = -4 \frac{dy}{y}$$

and, for $y = 200$ and $dy = 10$, the relative error in x is

$$\frac{dx}{x} = \frac{40}{200} = -\frac{1}{5}$$

and the percentage error is -20 . The negative sign of the error indicates that the errors in estimating x and y are opposite in sign.

PROBLEMS

Use differentials to find the approximate value of each of the following.

1. $\sqrt[3]{1010}$
2. $\sqrt[4]{15}$
3. $\sqrt{66}$
4. $\sqrt[3]{34}$

Find dy and Δy for each of the following.

5. $y = x^4 - x^2/2$ for $x = 2$, $\Delta x = 0.1$
6. $y = 12.8/y$ for $x = 10$, $\Delta x = 0.24$
7. $y = (x + 1)^3$ for $x = -3$, $\Delta x = -0.003$
8. $y = \sqrt{x}$ for $x = 4$, $\Delta x = 0.04$
9. A container is made in the form of a 10-cm cube to hold 1 liter (1000 cm^3). How accurately must the inner edge be made so that the volume will be correct to within 3 cm^3 ?
10. Using differentials, find the allowable percentage error in the diameter of a circle if the area is to be correct to within 4 percent.
11. Show that the relative error in the n th power of a measurement is approximately n times the relative error in the measurement.
12. Show that the relative error in the n th root of a measurement is approximately $1/n$ times the relative error in the measurement.

Answers to Odd-Numbered Problems

1. 10.033
3. 8.125
5. $dy = 3$; $\Delta y = 3.2431$
7. $dy = -0.036$; $\Delta y = -0.036027$
9. error must not exceed 0.01 cm

2.8 Higher-Order Derivatives

In some problems it is useful to differentiate a function more than once. The result of two or more successive differentiations of a function is a higher-order derivative.

The derivative of $y = f(x)$ with respect to x is, in general, a function of x and may be differentiated with respect to x . The derivative of the first derivative is the *second derivative*; its derivative is the *third derivative*, and so on.

In general, the n th derivative of a function $y = f(x)$ is obtained by differentiating n times. The result of this differentiation—that is, the n th derivative—is denoted by

$$\frac{d^n y}{dx^n} \quad f^{(n)}(x) \quad y^{(n)} \quad \frac{d^n}{dx^n}(y) \quad D_x^n y \quad D_x^n(y)$$

EXAMPLE

If $y = x^3 - 3x^2 + 2$,

$$\frac{dy}{dx} = 3x^2 - 6x$$

$$\frac{d^2 y}{dx^2} = 6x - 6$$

$$\frac{d^3 y}{dx^3} = 6$$

$$\frac{d^4 y}{dx^4} = 0$$

and higher-order derivatives are also zero.

EXAMPLE

If $x = \sqrt{t^2 + 1}$,

$$\begin{aligned}\frac{dx}{dt} &= \frac{1}{2}(t^2 + 1)^{-1/2}(2t) \\ &= t(t^2 + 1)^{-1/2}\end{aligned}$$

$$\begin{aligned}\frac{d^2x}{dt^2} &= (t^2 + 1)^{-1/2} + \left(-\frac{1}{2}\right)t(t^2 + 1)^{-3/2}(2t) \\ &= (t^2 + 1)^{-3/2}[t^2 + 1 - t^2] \\ &= (t^2 + 1)^{-3/2}\end{aligned}$$

$$\begin{aligned}\frac{d^3x}{dt^3} &= -\frac{3}{2}(t^2 + 1)^{-5/2}(2t) \\ &= -3t(t^2 + 1)^{-5/2}\end{aligned}$$

$$\begin{aligned}\frac{d^4x}{dt^4} &= -3(t^2 + 1)^{-5/2} - \frac{5}{2}(-3t)(t^2 + 1)^{-7/2}(2t) \\ &= (t^2 + 1)^{-7/2}(-3t^2 - 3 + 15t^2) \\ &= 3(4t^2 - 1)(t^2 + 1)^{-7/2}\end{aligned}$$

and so forth for higher-order derivatives.

EXAMPLE

If $y = \sin x$,

$$\frac{dy}{dx} = \cos x$$

$$\frac{d^2y}{dx^2} = -\sin x$$

$$\frac{d^3y}{dx^3} = -\cos x$$

$$\frac{d^4y}{dx^4} = \sin x$$

and so forth for higher-order derivatives.

Just as the derivative of the function $y = f(x)$ with respect to x represents the rate of change in y as x changes, the second derivative of $y = f(x)$ with respect to x represents the rate of change in the first derivative $y' = f'(x)$ as x changes.

In general, the n th derivative with respect to x of a function $y = f(x)$ represents the rate of change in the $(n - 1)$ st derivative of $y = f(x)$ as x changes.

Higher-order derivatives are very important for certain theoretical problems in mathematics and statistics; however, in most applied work, derivatives of order higher than the second are not frequently encountered.

Second-order derivatives are useful in graphical representation of functions, as discussed in later sections, and are also useful whenever the rate of change of the quantity represented by the first derivative is of interest. For example, the rate of change of velocity is acceleration and the rate of change of marginal cost indicates

whether marginal cost is increasing, decreasing, or constant under varying conditions of production.

EXAMPLE

If the position of a moving body at time t is represented by

$$s = (2t + 3)^2$$

find its velocity, $v = \frac{ds}{dt}$, and its acceleration, $a = \frac{dv}{dt} = \frac{d^2s}{dt^2}$.

$$v = \frac{ds}{dt} = 2(2t + 3)(2)$$

$$= 8t + 12$$

$$a = \frac{dv}{dt} = 8$$

EXAMPLE

An object projected vertically upward with a speed of 160 feet/second reaches an elevation of

$$s = 160t - 16t^2$$

at the end of t seconds.

- How high does it rise?
- How fast is it traveling when it reaches an elevation of 256 feet going up and again when it reaches this elevation coming down?
- What is its acceleration?

(a) $v = \frac{ds}{dt} = 160 - 32t$ and $v = 0$ when $t = 5$. Thus its maximum height is

$$s = 160(5) - 16(5^2)$$

$$= 800 - 400$$

$$= 400 \text{ ft}$$

(b) If $s = 256$,

$$256 = 160t - 16t^2$$

$$t^2 - 10t + 16 = 0$$

$$(t - 8)(t - 2) = 0$$

$$t = 8, 2$$

The object is thus at an elevation of 256 feet going up after 2 seconds and coming down after 8 seconds.

If $t = 2$, $v = \frac{ds}{dt} = 160 - 32t = 160 - 64 = 96 \text{ ft/sec}$

If $t = 8$, $v = \frac{ds}{dt} = 160 - 32t = 160 - 256 = -96 \text{ ft/sec}$

(c) $a = \frac{dv}{dt} = -32 \text{ ft/sec}^2$

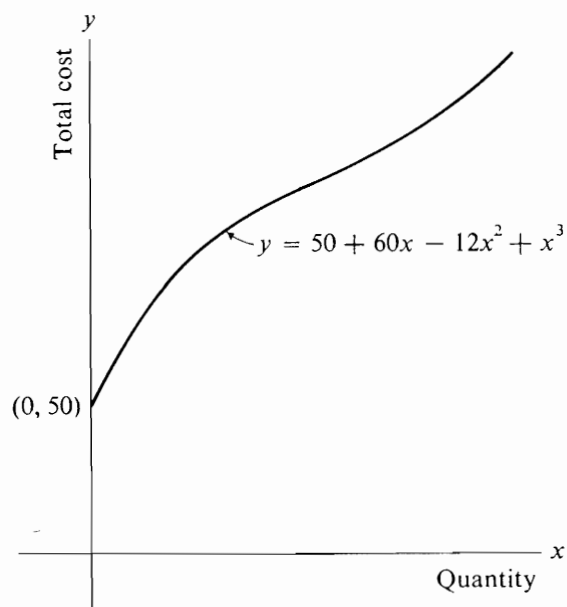
EXAMPLE

Total cost is represented by

$$y = 50 + 60x - 12x^2 + x^3$$

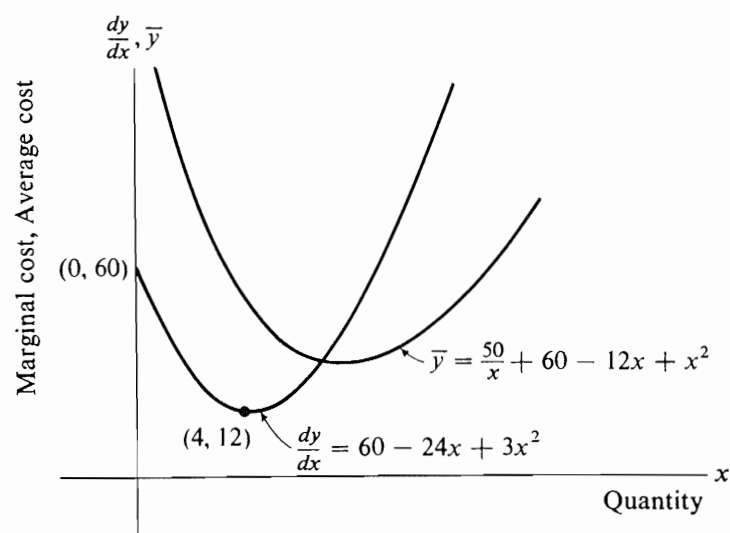
(see Figure 2.21).

Figure 2.21



where y is total cost and x is output. Average cost is $y/x = (50/x) + 60 - 12x + x^2$ and marginal cost is $\frac{dy}{dx} = 60 - 24x + 3x^2$ (see Figure 2.22).

Figure 2.22



$$\frac{d^2y}{dx^2} = -24 + 6x$$

$$= 0 \quad \text{when } x = 4$$

$$< 0 \quad \text{when } x < 4$$

$$> 0 \quad \text{when } x > 4$$

Thus marginal cost decreases for $x < 4$, increases for $x > 4$, and is constant for $x = 4$.

EXAMPLE

Total cost is represented by

$$y = 200 + 1000x - 24x^2 + 4x^3 + x^4$$

(see Figure 2.23).

where y is total cost and x is output. Average cost is

$$\frac{y}{x} = \frac{200}{x} + 1000 - 24x + 4x^2 + x^3$$

and marginal cost is

$$\frac{dy}{dx} = 1000 - 48x + 12x^2 + 4x^3$$

(see Figure 2.24).

$$\frac{d^2y}{dx^2} = -48 + 24x + 12x^2$$

$$= 0 \quad \text{when } x^2 + 2x - 4 = 0$$

Figure 2.23

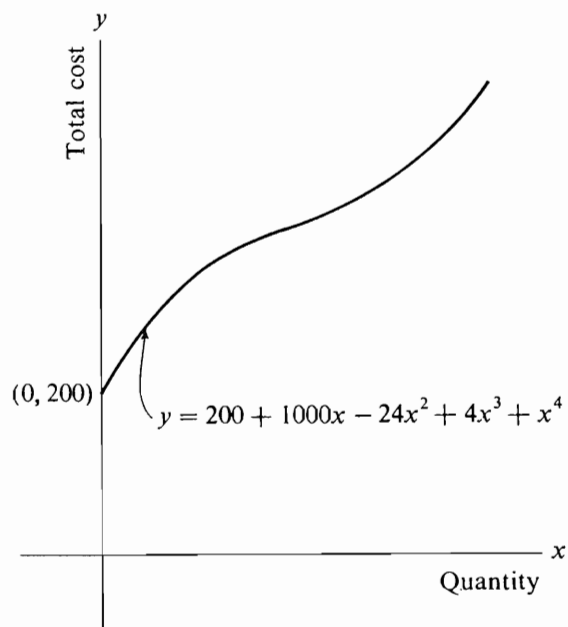
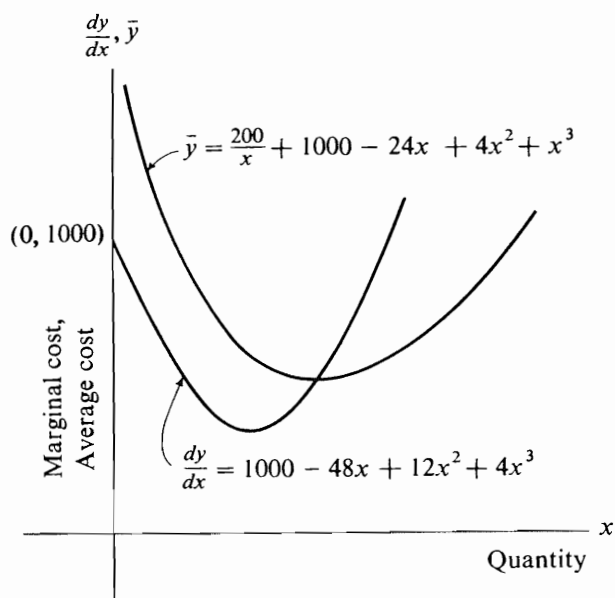


Figure 2.24



$$x = \frac{-2 \pm \sqrt{4 + 16}}{2}$$

$$= \frac{-2 \pm \sqrt{20}}{2} = -1 \pm \sqrt{5}$$

Appropriate value is $x = -1 + \sqrt{5}$ (approximately $x = 1.236$).

$$\frac{d^2y}{dx^2} = 0 \quad \text{when } x = -1 + \sqrt{5}$$

$$< 0 \quad \text{when } x < -1 + \sqrt{5}$$

$$> 0 \quad \text{when } x > -1 + \sqrt{5}$$

Thus marginal cost decreases for $x < -1 + \sqrt{5}$, increases for $x > -1 + \sqrt{5}$, and is constant for $x = -1 + \sqrt{5}$.

PROBLEMS

Find the first and second derivatives of each of the following functions.

1. $y = e^{3x}$
2. $y = x \ln x$
3. $y = \log \left(\frac{3}{x} \right)$
4. $y = x^2 e^x$
5. $y = e^{\ln(x^3 - 3)}$
6. $y = \left(\frac{x - 5}{x + 1} \right)^{-1}$

Answers to Odd-Numbered Problems

1. $3e^{3x}, 9e^{3x}$
3. $-\frac{1}{x} \log e, \frac{1}{x^2} \log e$
5. $3x^2, 6x$

2.9 Implicit Differentiation

Functions of the form $y = f(x)$ express y explicitly in terms of x and can be differentiated according to the rules for the types of functions involved. However, as noted in preceding sections, some equations involving x and y , of the form $f(x, y) = 0$, do not give y explicitly in terms of x and cannot conveniently be written so that they do. Such equations define y as a function of x in the sense that for each value of x there is a corresponding value of y which satisfies the equation; the equation is said to determine y as an *implicit function* of x . It is possible to calculate $\frac{dy}{dx}$ from such an equation by the method of *implicit differentiation*. In this method, y is treated as an unknown but differentiable function of x and the rules for finding derivatives are applied. In general, this method gives an expression for $\frac{dy}{dx}$ in terms of both x and y . Higher-order derivatives can also be obtained by implicit differentiation. The procedure for implicit differentiation is discussed and illustrated below.

Procedure for Implicit Differentiation

When y is defined as an implicit function of x by an equation $f(x, y) = 0$, the derivative of y with respect to x is obtained by differentiating the equation $f(x, y) = 0$ term by term, regarding y as a function of x , and then solving the resulting equation for the derivative $\frac{dy}{dx}$.

EXAMPLE

Find $\frac{dy}{dx}$ for the equation $xy^2 - x^2 + y = 0$. Differentiating with respect to x ,

$$2xy \frac{dy}{dx} + y^2 - 2x + \frac{dy}{dx} = 0$$

$$(2xy + 1) \frac{dy}{dx} = 2x - y^2$$

$$\frac{dy}{dx} = \frac{2x - y^2}{2xy + 1}$$

NOTE: In this example, xy^2 is the product of two functions of x , x and y^2 , and is differentiated using the product rule

$$\frac{d}{dx}(uv) = u \frac{dv}{dx} + v \frac{du}{dx}$$

where $u = x$, $v = y^2$, $\frac{du}{dx} = \frac{dx}{dx} = 1$, and $\frac{dv}{dx} = 2y \frac{dy}{dx}$. Thus $\frac{d}{dx}(xy^2) = 2xy \frac{dy}{dx} + y^2$. The derivative of $-x^2$ is $-2x$, as in differentiation of explicit functions (note again that $\frac{dx}{dx} = 1$). The derivative of y is $\frac{dy}{dx}$, again as in differentiation of explicit functions.

EXAMPLE

Find $\frac{dy}{dx}$ for the equation $x^3 + y^3 - 3axy = 0$. Differentiating with respect to x ,

$$3x^2 + 3y^2 \frac{dy}{dx} - 3ax \frac{dy}{dx} - 3ay = 0$$

$$(y^2 - ax) \frac{dy}{dx} = ay - x^2$$

$$\frac{dy}{dx} = \frac{ay - x^2}{y^2 - ax}$$

EXAMPLE

Find $\frac{dy}{dx}$ for the equation $\sin x \cos y - \tan y = 0$. Differentiating with respect to x ,

$$\cos x \cos y - \sin x \sin y \frac{dy}{dx} - \sec^2 y \frac{dy}{dx} = 0$$

$$\frac{dy}{dx} = \frac{\cos x \cos y}{\sin x \sin y + \sec^2 y}$$

EXAMPLE

Find $\frac{dy}{dx}$ for the equation $x^2 - y^2 = 1$. Differentiating with respect to x ,

$$2x - 2y \frac{dy}{dx} = 0$$

$$\frac{dy}{dx} = \frac{x}{y}$$

Differentiating $\frac{dy}{dx}$ with respect to x ,

$$\frac{d^2y}{dx^2} = \frac{y - x \frac{dy}{dx}}{y^2}$$

Substituting $\frac{dy}{dx} = \frac{x}{y}$,

$$\begin{aligned} \frac{d^2y}{dx^2} &= \frac{y - x \frac{x}{y}}{y^2} \\ &= \frac{y^2 - x^2}{y^3} \end{aligned}$$

or, since $y^2 - x^2 = -1$,

$$\frac{d^2y}{dx^2} = -\frac{1}{y^3}$$

EXAMPLE

Find $\frac{dy}{dx}$ and $\frac{d^2y}{dx^2}$ for the equation $e^y - xe^x = 0$. Differentiating with respect to x ,

$$e^y \frac{dy}{dx} - e^x - xe^x = 0$$

$$\frac{dy}{dx} = \frac{e^x(x+1)}{e^y}$$

or, since $e^y = xe^x$,

$$\frac{dy}{dx} = \frac{x+1}{x}$$

and

$$\frac{d^2y}{dx^2} = \frac{x - (x+1)}{x^2}$$

$$= -\frac{1}{x^2}$$

EXAMPLE

$x^{1/2} + y^{1/2} = a^{1/2}$. Differentiating with respect to x ,

$$\frac{1}{2}x^{-1/2} + \frac{1}{2}y^{-1/2} \frac{dy}{dx} = 0$$

$$\frac{dy}{dx} = -\frac{y^{1/2}}{x^{1/2}}$$

Differentiating $\frac{dy}{dx}$ with respect to x ,

$$\frac{d^2y}{dx^2} = -\frac{\left(\frac{1}{2}y^{-1/2} \frac{dy}{dx}\right)x^{1/2} - y^{1/2}\left(\frac{1}{2}x^{-1/2}\right)}{x}$$

Substituting $\frac{dy}{dx} = -\frac{y^{1/2}}{x^{1/2}}$,

$$\frac{d^2y}{dx^2} = \frac{\frac{1}{2} + \frac{1}{2}x^{-1/2}y^{1/2}}{x}$$

$$= \frac{1}{2}x^{-1} + \frac{1}{2}x^{-3/2}y^{1/2}$$

$$= \frac{1}{2}x^{-3/2}(x^{1/2} + y^{1/2})$$

or, since $x^{1/2} + y^{1/2} = a^{1/2}$,

$$\frac{d^2y}{dx^2} = \frac{1}{2}a^{1/2}x^{-3/2}$$

PROBLEMS

Find a. $\frac{dy}{dx}$ and b. $\frac{d^2y}{dx^2}$ for each of the following functions by the method of implicit differentiation.

1. $x^2 + y^2 = 1$
2. $x^3 + y^3 = 1$
3. $x^{2/3} + y^{2/3} = 1$
4. $xy + y^2 = 1$
5. $y^2 = x^3$
6. $xy = a$
7. $x^2y^2 = b$
8. $x^2y^3 = c$
9. $x + y - xy = 2$
10. $x^2y^2 + xy = 1$
11. $x^3y^3 + x^2y^2 = a$
12. $x + y + xy + y^2 = b$
13. $x^3y^2 = a$
14. $x^2 + xy = a^2$
15. $xy^2 + y^2 = a$
16. $xy + y^3 = b$
17. $(x + y)^2 + (x + y)^3 = a^2$
18. $x^2 + y^2 = ab$

Find $\frac{dy}{dx}$ for each of the following functions by the method of implicit differentiation.

19. $y^2 = \frac{x-1}{x+1}$
20. $x^3 - xy + y^3 = 1$
21. $x^2 = \frac{x-y}{x+y}$
22. $y = x(x^2 + 1)^{-1/2}$
23. $y = x^{1/2} + x^{1/3} + x^{1/4}$
24. $y^2 = \frac{x^2 - 1}{x^2 + 1}$
25. $(x + y)^3 + (x - y)^3 = x^4 + y^4$
26. $y = (x + 5)^4(x^2 - 2)^3$
27. $\frac{1}{y} + \frac{1}{x} = 1$
28. $y = (x^2 + 3)^{1/3}x^{-1}$
29. $\cos^2 x + \tan y \sin y = 0$
30. $\csc x - \sec y + \tan y + \cot x = 0$
31. $xye^x + e^y = 0$
32. $ye^x = 10 + ye^y$
33. $x^5 + 4xy^3 - 3y^5 = 2$
34. $x^2 + xy + y^2 = 1$

Answers to Odd-Numbered Problems

1. a. $-\frac{x}{y}$ b. $-\frac{1}{y^3}$
3. a. $-x^{-1/3}y^{1/3}$ b. $\frac{1}{3}x^{-4/3}y^{-1/3}$
5. a. $\frac{3x^2}{2y}$ b. $\frac{3x}{4y}$
7. a. $-\frac{y}{x}$ b. $\frac{2y}{x^2}$
9. a. $\frac{y-1}{1-x}$ b. $\frac{2(y-1)}{(1-x)^2}$
11. a. $-\frac{y}{x}$ b. $\frac{2y}{x^2}$
13. a. $-\frac{3y}{2x}$ b. $\frac{15y}{4x^2}$
15. a. $\frac{-y}{2(x+1)}$ b. $\frac{3y}{4(x+1)^2}$
17. a. -1 b. 0
19. $\frac{1}{y(x+1)^2}$
21. $\frac{y}{x} - (x+y)^2$
23. $\frac{1}{2}x^{-1/2} + \frac{1}{3}x^{-2/3} + \frac{1}{4}x^{-3/4}$
25. $\frac{2x^3 - 3x^2 - 3y^2}{6xy - 2y^3}$
27. $-\frac{y^2}{x^2}$
29. $\frac{2 \sin x \cos x}{\sin y(1 + \sec^2 y)}$
31. $-\frac{y(1+x)}{x(1-y)}$
33. $\frac{5x^4 + 4y^3}{15y^4 - 12xy^2}$

2.10 Differentiability and Continuity

The relationships between differentiability and continuity are discussed in this section and smooth functions are defined and illustrated.

It can be shown that if the function $y = f(x)$ has a finite derivative

$$f'(c) = \lim_{\Delta x \rightarrow 0} \frac{f(c + \Delta x) - f(c)}{\Delta x}$$

at $x = c$, then $f(x)$ is continuous at $x = c$.

Continuity at a point does not, however, imply the existence of a derivative at that point. Thus

Differentiability $\not\Rightarrow$ Continuity

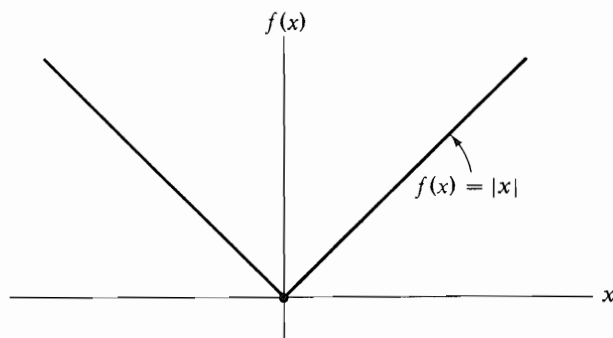
EXAMPLE

$f(x) = |x|$ is continuous at $x = 0$, but does *not* have a derivative at $x = 0$, since

$$\begin{aligned} f'(x) &= 1 && \text{for } x > 0 \\ &= -1 && \text{for } x < 0 \end{aligned}$$

(see Figure 2.25).

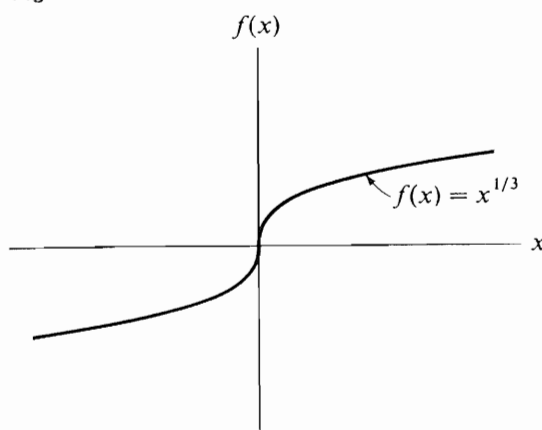
Figure 2.25



EXAMPLE

$f(x) = x^{1/3}$ is continuous at $x = 0$, but does *not* have a derivative at $x = 0$, since $f'(x) = \frac{1}{3}x^{-2/3}$ is not defined for $x = 0$ (see Figure 2.26).

Figure 2.26



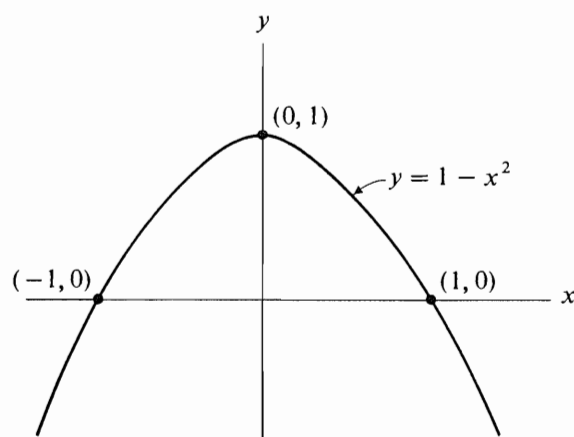
Smooth Functions

A function $f(x)$ is said to be a *smooth function* if both $f(x)$ and its first derivative are continuous. The graph of a smooth function is said to be a *smooth curve*; such a curve is not only continuous and unbroken but also has a continuously turning tangent. As shown in the following sections, a smooth function has some useful properties not possessed by a continuous curve whose first derivative is discontinuous at certain points.

EXAMPLE

$y = 1 - x^2$ is a smooth function, since y and its derivative $\frac{dy}{dx} = -2x$ are everywhere continuous (see Figure 2.27).

Figure 2.27



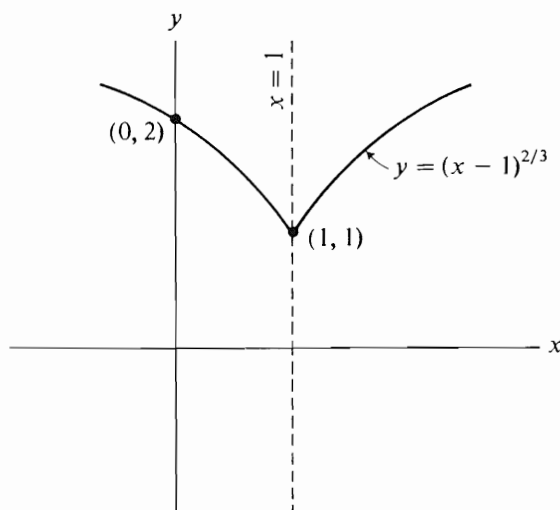
EXAMPLE

$y = (x - 1)^{2/3} + 1$ is not a smooth function at $x = 1$, since its derivative

$$\frac{dy}{dx} = \frac{2}{3\sqrt[3]{x-1}}$$

is discontinuous (becomes infinite) at $x = 1$ (see Figure 2.28).

Figure 2.28



2.11 Applications of Derivatives

In this section the information concerning a function or curve which can be obtained from its first and second derivatives is discussed. Such information is useful not only for graphing but also for determining the appropriateness of equations for representing various relationships in business, economics, and the behavioral sciences.

THE FIRST DERIVATIVE

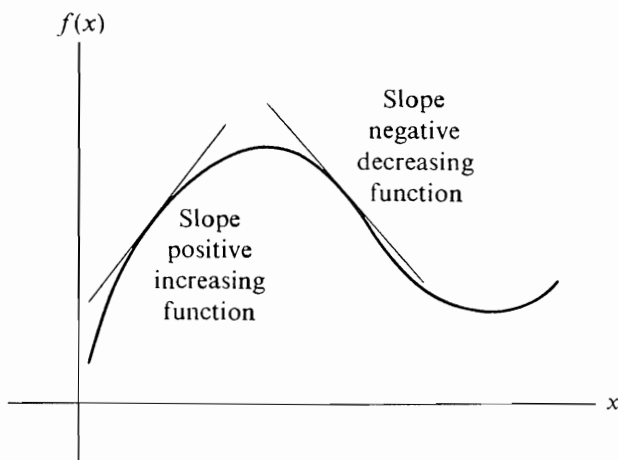
The first derivative of a function can be used to determine where the function is increasing and where it is decreasing. The first derivative can also be used to locate maximum and minimum values of the function (stationary points), if there are any.

Increasing and Decreasing Functions

Consider the function $y = f(x)$ at the point $x = a$. The first derivative of y with respect to x , $y' = f'(x)$ is the slope at the point x of the curve representing the function $y = f(x)$; in particular, $f'(a)$ is the slope of the curve $y = f(x)$ at the point $x = a$.

If the first derivative $f'(a)$ is positive, $y = f(x)$ is an *increasing function* of x at $x = a$; that is, $y = f(x)$ increases as x increases past $x = a$; if the first derivative $f'(a)$ is negative, $y = f(x)$ is a *decreasing function* of x at $x = a$; that is, $y = f(x)$ decreases as x increases past $x = a$. See Figure 2.29.

Figure 2.29

**EXAMPLE**

$f(x) = 2x^2 + 10$, $f'(x) = 4x$. Since $f'(x) > 0$ for $x > 0$ and $f'(x) < 0$ for $x < 0$, $f(x)$ is an increasing function of x for positive x and a decreasing function of x for negative x . Note that $f'(x) = 0$ for $x = 0$.

EXAMPLE

$f(x) = 3x^3$, $f'(x) = 9x^2$. Since $f'(x)$ is nonnegative for all values of x , $f(x)$ is an increasing function of x for all values of x .

EXAMPLE

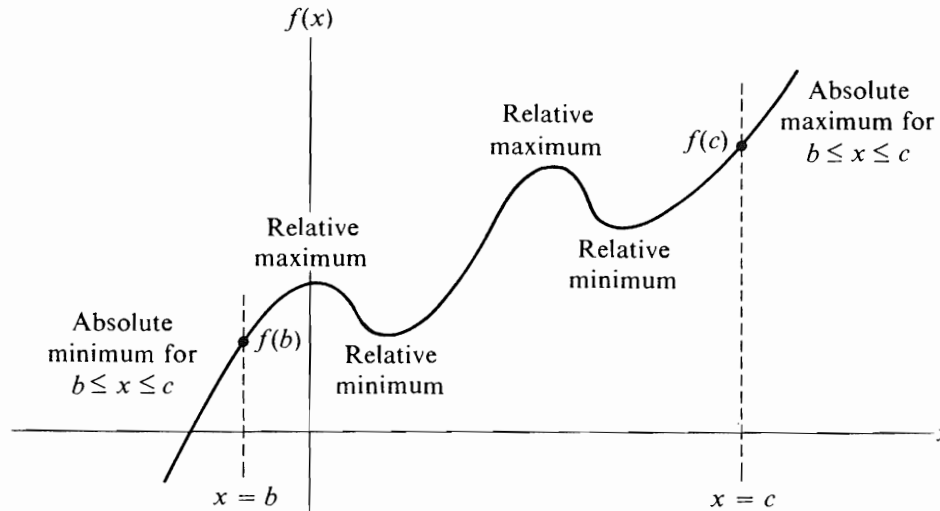
$f(x) = x^3 + 6x^2 + 15$, $f'(x) = 3x^2 + 12x = 3x(x + 4)$. Since $f'(x) > 0$ for $x > 0$ or $x < -4$ and $f'(x) < 0$ for $-4 < x < 0$, $f(x)$ is a decreasing function of x for the interval $-4 < x < 0$ and an increasing function of x for all $x < -4$ and $x > 0$. Note that $f'(x) = 0$ for $x = 0$.

In the examples above, the values of x for which $f'(x) = 0$ are not discussed; such values are of particular interest, as discussed below.

Relative (Local) Maxima and Minima

A function $y = f(x)$ is said to have a *relative maximum* or *local maximum* at $x = a$ if $f(a)$ is greater than any value of $f(x)$ for x in an interval around a ; a function $y = f(x)$ is said to have a *relative minimum* or *local minimum* at $x = a$ if $f(a)$ is smaller than any value of $f(x)$ for x in an interval around a . Note that a relative maximum or minimum of a function is its maximum or minimum for a given interval; the (absolute) maximum or minimum of the function over a larger interval can occur at an endpoint of the interval, rather than at any relative maximum or minimum. Also note

Figure 2.30



that it is possible for a relative maximum value of a function to be less than a relative minimum value of the function (see Figure 2.30).

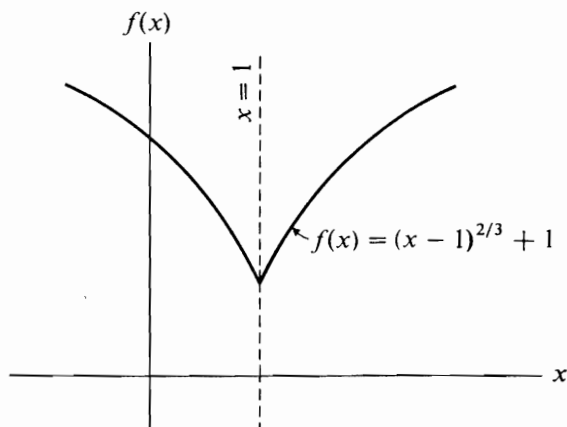
If a function $f(x)$ has a relative maximum or minimum at a value $x = a$ for which its first derivative is continuous, then $f'(a) = 0$. Values of x for which $f'(x)$ is discontinuous must be considered separately, as discussed below.

CAUTIONS

1. A relative maximum or minimum at $x = a$ implies $f'(a) = 0$ only if $f(x)$ and $f'(x)$ are continuous at $x = a$.
2. $f'(a) = 0$ does *not* imply a relative maximum or minimum at $x = a$ even if $f(x)$ and $f'(x)$ are continuous at $x = a$; that is, if $f(x)$ and $f'(x)$ are continuous at $x = a$, $f'(a) = 0$ is a necessary but not a sufficient condition for a relative maximum or minimum at $x = a$.

EXAMPLE (CAUTION 1)

Figure 2.31



If $f(x) = (x - 1)^{2/3} + 1$, then

$$f'(x) = \frac{2}{3(x - 1)^{1/3}}$$

and $f'(x)$ is (infinitely) discontinuous at $x = 1$. Thus, although the function has a relative minimum at $x = 1$, $f'(1) \neq 0$ (see Figure 2.31).

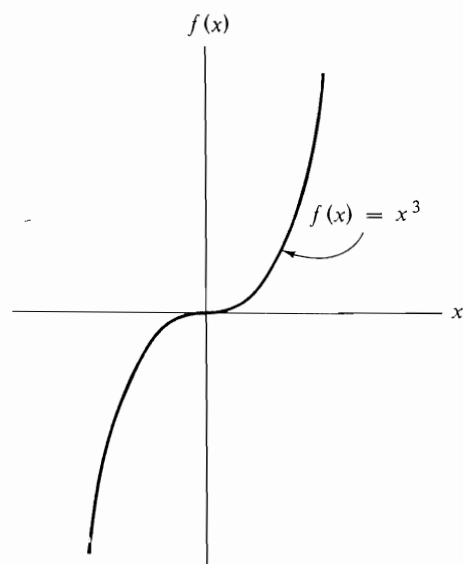
EXAMPLE (CAUTION 2)

If $f(x) = x^3$, then

$$f'(x) = 3x^2$$

and $f'(x) = 0$ for $x = 0$. However, the function $f(x) = x^3$ does not have a relative maximum or minimum at $x = 0$ (see Figure 2.32).

Figure 2.32

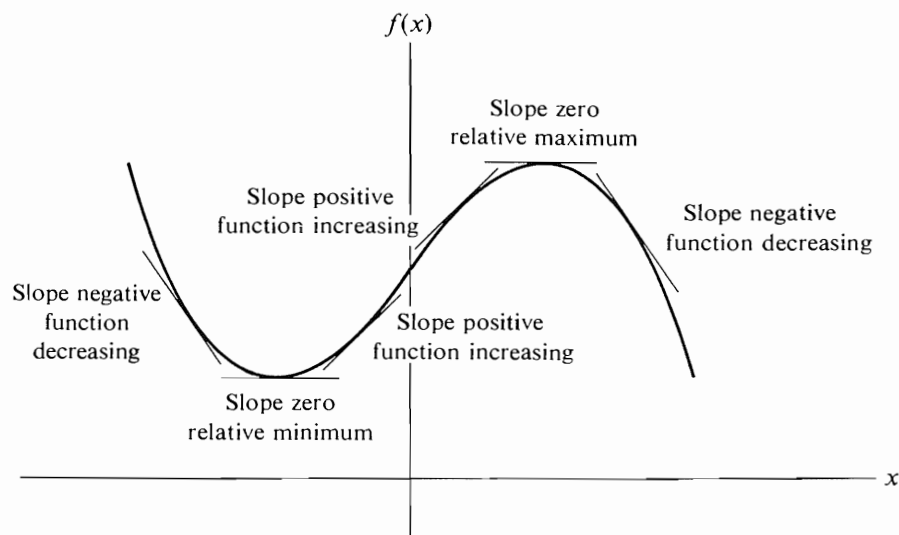


Consider a function $f(x)$ at a value $x = a$ for which $f(x)$ and $f'(x)$ are continuous. It is evident geometrically that if $f(a)$ is a relative maximum of $f(x)$, then the slope $f'(x)$ of $f(x)$ changes from positive to negative as x passes through the point $x = a$; similarly, if $f(a)$ is a relative minimum of $f(x)$, then the slope $f'(x)$ of $f(x)$ changes from negative to positive as x passes through the point $x = a$ (see Figure 2.33). The corresponding algebraic argument is based on the fact that an increasing function has a positive slope and a decreasing function has a negative slope.

Thus, in order to determine the relative maximum and minimum values, if any, of a function $y = f(x)$,

1. Solve the equation $f'(x) = 0$ to obtain its roots (sometimes called critical values).
2. For each root a , determine whether $f'(x)$ changes sign as x increases through a :

Figure 2.33



$f'(x)$ changes from $+$ to $-$ at $x = a \Rightarrow$ relative maximum at $x = a$

$f'(x)$ changes from $-$ to $+$ at $x = a \Rightarrow$ relative minimum at $x = a$

$f'(x)$ does not change sign at $x = a \Rightarrow$ no relative maximum or minimum at $x = a$

NOTE: The above procedure locates all relative maxima and minima that occur at values of x for which $f(x)$ and $f'(x)$ are continuous. Relative maxima and minima which occur at values of x for which $f'(x)$ is discontinuous are discussed below.

EXAMPLE

Find the relative maxima and minima (if any) of the function $y = 2x^3 - 3x^2 - 12x + 13$.

$$\frac{dy}{dx} = 6x^2 - 6x - 12$$

$$= 0 \quad \text{if } x^2 - x - 2 = 0$$

$$(x - 2)(x + 1) = 0$$

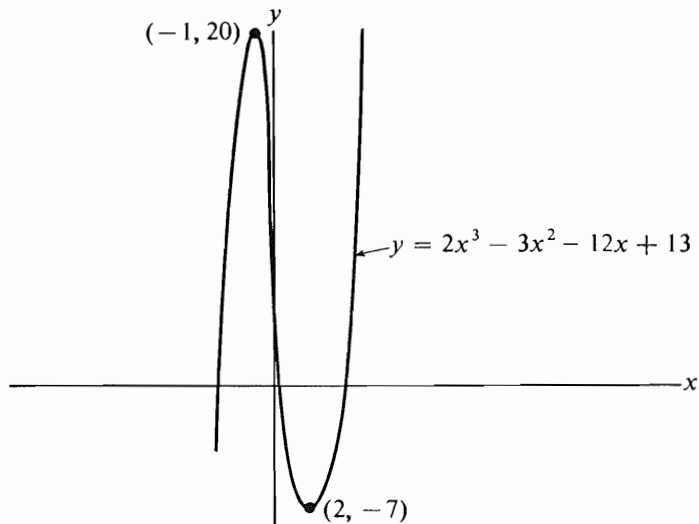
$$x = 2, -1$$

$$\left. \begin{array}{l} \text{If } -1 < x < 2, \frac{dy}{dx} < 0 \\ \text{If } x > 2, \frac{dy}{dx} > 0 \end{array} \right\} \text{so minimum at } x = 2$$

$$\left. \begin{array}{l} \text{If } x < -1, \frac{dy}{dx} > 0 \\ \text{If } -1 < x < 2, \frac{dy}{dx} < 0 \end{array} \right\} \text{so maximum at } x = -1$$

See Figure 2.34.

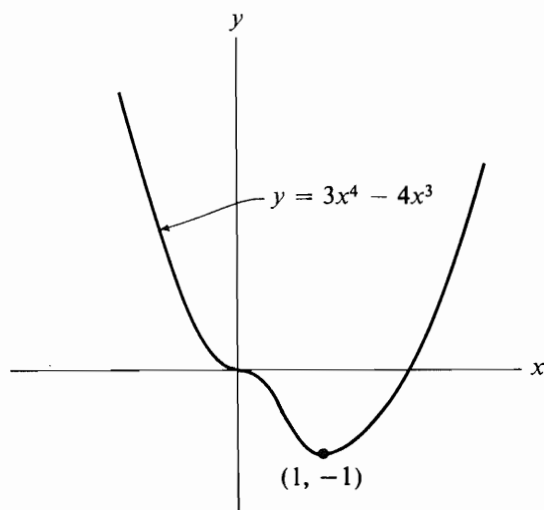
Figure 2.34

**EXAMPLE**

Find the relative maxima and minima (if any) of the function $y = 3x^4 - 4x^3$.

$$\begin{aligned}\frac{dy}{dx} &= 12x^3 - 12x^2 \\ &= 0 \quad \text{if } x^3 - x^2 = 0 \\ &\quad x^2(x - 1) = 0 \\ &\quad x = 0, 1\end{aligned}$$

Figure 2.35



$$\begin{array}{l}
 \text{If } x < 0, \frac{dy}{dx} < 0 \\
 \text{If } 0 < x < 1, \frac{dy}{dx} < 0 \\
 \text{If } 0 < x < 1, \frac{dy}{dx} < 0 \\
 \text{If } x > 1, \frac{dy}{dx} > 0
 \end{array}
 \left. \vphantom{\begin{array}{l} \text{If } x < 0, \frac{dy}{dx} < 0 \\ \text{If } 0 < x < 1, \frac{dy}{dx} < 0 \\ \text{If } 0 < x < 1, \frac{dy}{dx} < 0 \\ \text{If } x > 1, \frac{dy}{dx} > 0 \end{array}} \right\}
 \begin{array}{l}
 \text{so no maximum or minimum at } x = 0 \\
 \\
 \text{so minimum at } x = 1
 \end{array}$$

See Figure 2.35.

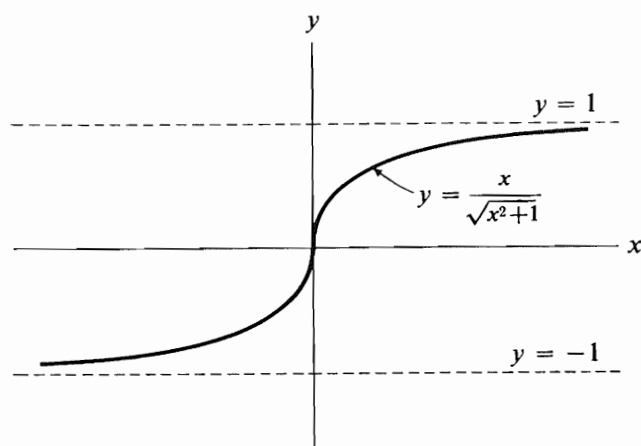
EXAMPLE

Find the relative maxima and minima (if any) of the function $y = x/\sqrt{x^2 + 1}$.

$$\begin{aligned}
 \frac{dy}{dx} &= \frac{(x^2 + 1)^{1/2} - \frac{1}{2}x(x^2 + 1)^{-1/2}(2x)}{(x^2 + 1)} \\
 &= \frac{(x^2 + 1)^{-1/2}(x^2 + 1 - x^2)}{(x^2 + 1)} \\
 &= \frac{1}{(x^2 + 1)^{3/2}} \\
 &\neq 0
 \end{aligned}$$

for all values of x , so there is no maximum or minimum (see Figure 2.36).

Figure 2.36



In each of the examples above, the first derivative is continuous for all values of x . Thus relative maxima and minima can occur only for values $x = a$ for which $f'(a) = 0$. As noted above, relative maxima or minima may also occur for values $x = a$ for which $f'(a)$ is discontinuous. That is, if $f(x)$ is continuous at $x = a$ but its

first derivative is discontinuous (finitely or infinitely) at $x = a$, then $f(a)$ may be a relative maximum or minimum even though $f'(a) \neq 0$. Again, a change in the sign of the first derivative as x passes through a is necessary for the existence of a relative maximum or minimum at $x = a$.

Thus, in order to determine relative maximum and minimum values (if any) of a function $y = f(x)$ at points where $f(x)$ is continuous but $f'(x)$ is discontinuous,

1. Determine the values of x for which $f'(x)$ is finitely or infinitely discontinuous and $f(x)$ is continuous.
2. For each value a such that $f(x)$ is continuous at $x = a$ and $f'(x)$ is discontinuous at $x = a$, determine whether $f'(x)$ changes sign as x increases through a :

$f'(x)$ changes from $+$ to $-$ at $x = a \Rightarrow$ relative maximum at $x = a$

$f'(x)$ changes from $-$ to $+$ at $x = a \Rightarrow$ relative minimum at $x = a$

$f'(x)$ does not change sign at $x = a \Rightarrow$ no relative maximum or minimum at $x = a$

EXAMPLE

Find the relative maxima and minima (if any) of the function $y = x^{2/3}$.

$$\frac{dy}{dx} = \frac{2}{3}x^{-1/3}$$

$\neq 0$ for all values of x

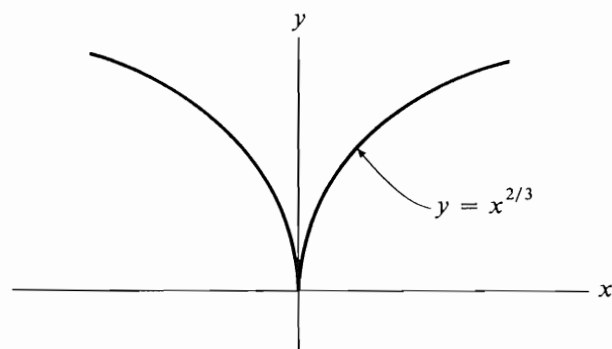
$$\frac{dy}{dx} \rightarrow -\infty \text{ as } x \rightarrow 0^- \quad \text{and} \quad \frac{dy}{dx} \rightarrow \infty \text{ as } x \rightarrow 0^+$$

so $\frac{dy}{dx}$ has an infinite discontinuity at $x = 0$.

$$\left. \begin{array}{l} \text{If } x < 0, \frac{dy}{dx} < 0 \\ \text{If } x > 0, \frac{dy}{dx} > 0 \end{array} \right\} \text{so minimum at } x = 0$$

See Figure 2.37.

Figure 2.37



EXAMPLE

Find the relative maxima and minima (if any) of the function $y = (1 - x)^{2/3}(2 + x)^{1/3}$.

$$\begin{aligned}\frac{dy}{dx} &= -\frac{2}{3}(1 - x)^{-1/3}(2 + x)^{1/3} + (1 - x)^{2/3}\left(\frac{1}{3}\right)(2 + x)^{-2/3} \\ &= \frac{1}{3}(1 - x)^{-1/3}(2 + x)^{-2/3}[-2(2 + x) + (1 - x)] \\ &= \frac{-3 - 3x}{3(1 - x)^{1/3}(2 + x)^{2/3}} \\ &= \frac{-(x + 1)}{(1 - x)^{1/3}(2 + x)^{2/3}} \\ &= 0 \quad \text{if } x = -1\end{aligned}$$

$$\frac{dy}{dx} \rightarrow -\infty \text{ as } x \rightarrow 1^- \quad \text{and} \quad \frac{dy}{dx} \rightarrow \infty \text{ as } x \rightarrow 1^+$$

so $\frac{dy}{dx}$ has an infinite discontinuity at $x = 1$.

$$\frac{dy}{dx} \rightarrow -\infty \text{ as } x \rightarrow -2 \quad \text{and} \quad \frac{dy}{dx} \rightarrow \infty \text{ as } x \rightarrow -2$$

so $\frac{dy}{dx}$ has an infinite discontinuity at $x = -2$.

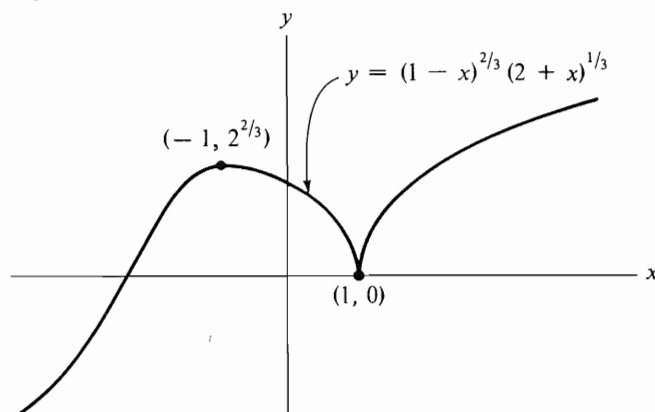
$$\left. \begin{array}{l} \text{If } -2 < x < -1, \frac{dy}{dx} > 0 \\ \text{If } -1 < x < 1, \frac{dy}{dx} < 0 \end{array} \right\} \text{so maximum at } x = -1$$

$$\left. \begin{array}{l} \text{If } -1 < x < 1, \frac{dy}{dx} < 0 \\ \text{If } x > 1, \frac{dy}{dx} > 0 \end{array} \right\} \text{so minimum at } x = 1$$

$$\left. \begin{array}{l} \text{If } x < -2, \frac{dy}{dx} > 0 \\ \text{If } -2 < x < -1, \frac{dy}{dx} > 0 \end{array} \right\} \text{so no maximum or minimum at } x = -2$$

See Figure 2.38.

Figure 2.38



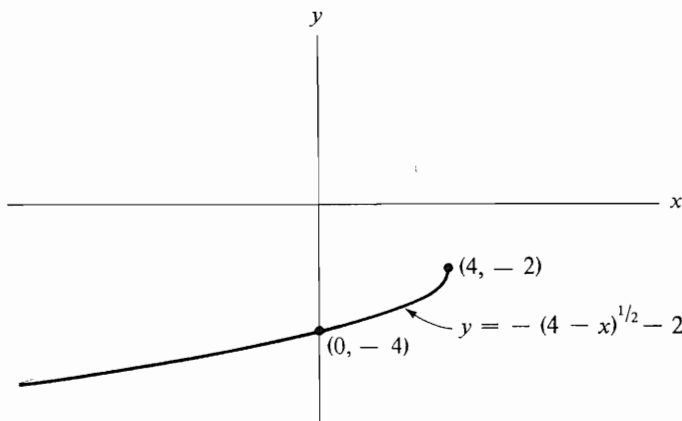
EXAMPLE

Find the relative maxima and minima (if any) of the function $y = -(4 - x)^{1/2} - 2$.

$$\begin{aligned}\frac{dy}{dx} &= \frac{1}{2}(4 - x)^{-1/2} \\ &= \frac{1}{2(4 - x)^{1/2}}\end{aligned}$$

$\frac{dy}{dx} \rightarrow \infty$ if $x \rightarrow 4^-$, so $\frac{dy}{dx}$ has an infinite discontinuity at $x = 4$. Note that y and $\frac{dy}{dx}$ are undefined if $x > 4$. Thus $y = -(4 - x)^{1/2} - 2$ has no relative maximum or minimum (see Figure 2.39).

Figure 2.39



The least (or greatest) value of a function in an interval—that is, its absolute minimum or maximum in an interval—may occur at an end point of the interval rather than at a relative minimum or maximum value.

EXAMPLE

Find the maximum and minimum values of the function $y = x^2 - 2x$ in the interval $2 \leq x \leq 4$.

$$\begin{aligned}\frac{dy}{dx} &= 2x - 2 \\ &= 0 \quad \text{if } x = 1 \\ &< 0 \quad \text{if } x < 1 \\ &> 0 \quad \text{if } x > 1\end{aligned} \quad \left. \vphantom{\begin{aligned} \frac{dy}{dx} &= 2x - 2 \\ &= 0 \quad \text{if } x = 1 \\ &< 0 \quad \text{if } x < 1 \\ &> 0 \quad \text{if } x > 1 \end{aligned}} \right\} \text{so minimum at } x = 1$$

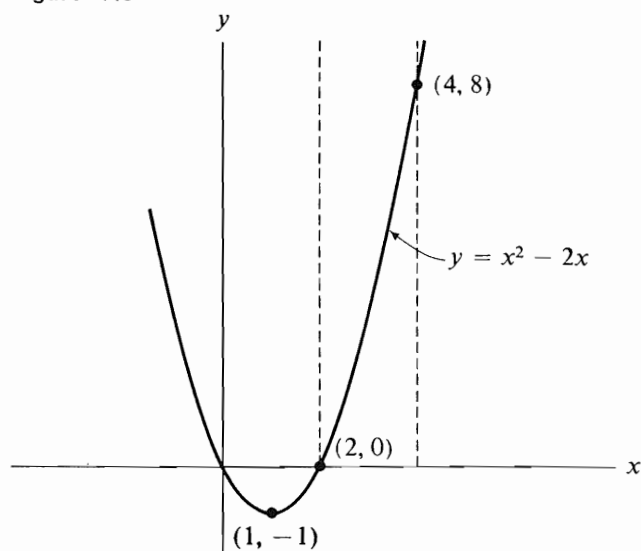
But $x = 1$ is outside the interval $2 \leq x \leq 4$.

$$\text{If } x = 2, y = 0$$

$$\text{If } x = 4, y = 8$$

Thus in the interval $2 \leq x \leq 4$, the least value of y occurs at the end point $x = 2$, and the greatest value occurs at the end point $x = 4$; at neither of these points is $\frac{dy}{dx}$ equal to zero (see Figure 2.40).

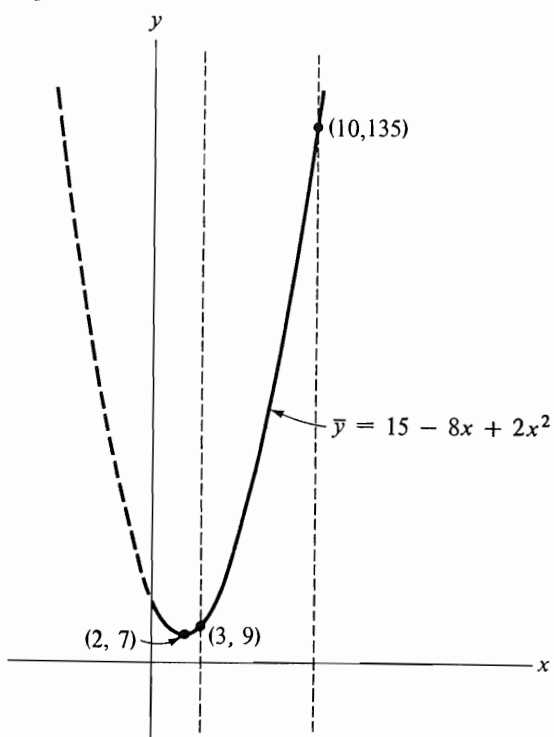
Figure 2.40



EXAMPLE

The total cost curve of a commodity is $y = 15x - 8x^2 + 2x^3$, where y represents total cost and x represents quantity. Suppose that market conditions indicate that between 3 and 10 units should be produced (that is, $3 \leq x \leq 10$). Determine the quantity in this interval for which average cost is a minimum.

Figure 2.41



$$\text{Average cost} = y/x = \bar{y} = 15 - 8x + 2x^2.$$

$$\frac{d\bar{y}}{dx} = -8 + 4x$$

$$= 0 \quad \text{if } x = 2$$

$$\begin{array}{ll} < 0 & \text{if } x < 2 \\ > 0 & \text{if } x > 2 \end{array} \left\} \text{so minimum at } x = 2$$

But $x = 2$ is outside the interval $3 \leq x \leq 10$.

$$\text{If } x = 3, \bar{y} = 9$$

$$\text{If } x = 10, \bar{y} = 135$$

Thus in the interval $3 \leq x \leq 10$, the least value of \bar{y} occurs at $x = 3$ and the greatest value occurs at $x = 10$; at neither of these points is $\frac{d\bar{y}}{dx}$ equal to zero. Thus, for between 3 and 10 units, average cost is minimum for 3 units (see Figure 2.41).

SUMMARY OF THE INFORMATION THAT CAN BE OBTAINED FROM THE FIRST DERIVATIVE

Consider the function $y = f(x)$ at the point $x = a$

$$f'(a) > 0 \Rightarrow f(x) \text{ is an increasing function of } x \text{ at } x = a$$

$$f'(a) < 0 \Rightarrow f(x) \text{ is a decreasing function of } x \text{ at } x = a$$

$$\text{If } \begin{cases} f(x) \text{ is continuous at } x = a \\ \text{and either} \\ f'(a) = 0 \\ \text{or} \\ f'(x) \text{ is discontinuous at } x = a \end{cases} \quad \text{then } \begin{cases} f'(x) \text{ changes from } + \text{ to } - \text{ at } \\ x = a \Rightarrow \text{relative maximum at } x = a \\ f'(x) \text{ changes from } - \text{ to } + \text{ at } \\ x = a \Rightarrow \text{relative minimum at } x = a \\ f'(x) \text{ does not change sign at } \\ x = a \Rightarrow \text{no relative maximum} \\ \text{or minimum at } x = a \end{cases}$$

Note that the least or greatest value of a function in an interval, that is, its absolute minimum or maximum in an interval, may occur at an end point of the interval rather than at a relative minimum or maximum value.

PROBLEMS

For each of the following functions, determine maxima and minima; sketch the curve representing each function.

$$1. y = 12 - 12x + x^3$$

$$2. y = \frac{x}{x+1}$$

$$3. y = \frac{x^3}{3} - \frac{x^2}{2} - 6x$$

$$4. y = x^4 - 32x + 48$$

$$5. y = \frac{x}{\sqrt{x^2 + 7}}$$

$$6. y = \frac{2x}{\sqrt{x^2 - 1}}$$

$$7. y = x^2 - 4x + 3$$

$$8. y = x\sqrt{1 - x^2}$$

$$9. y = x^3 - 3x^2 + 2$$

$$10. y = \frac{1}{x^2 + 4}$$

11. $y = \frac{x^2}{\sqrt{x^2 - 8}}$

12. $y = x^4 - 4x^3 + 12$

13. $y = \frac{3x^2}{\sqrt{x^2 + 3}}$

14. $y = \frac{1}{16 - x^2}$

15. $y = \frac{2}{3}x^3 - 4x^2 + 6x + 2$

16. $y = \frac{8x}{x^2 + 4}$

17. $y = \frac{x + 3}{x^2}$

18. $y = x^5 + 6$

19. $y = (x - 1)^{1/3}(x + 1)^{2/3}$

20. $y = \frac{1}{6}(x^3 - 6x^2 + 9x + 6)$

Answers to Odd-Numbered Problems

1. $\max(-2, 28)$

$\min(2, -4)$

3. $\max(-2, \frac{22}{3})$

$\min(3, -\frac{27}{2})$

5. no max

no min

7. no max

$\min(2, -1)$

9. $\max(0, 2)$

$\min(2, -2)$

11. no max

$\min(\pm 4, 4\sqrt{2})$

13. no max

$\min(0, 0)$

15. $\max(1, \frac{14}{3})$

$\min(3, 2)$

17. no max

$\min(-6, -\frac{1}{2})$

19. $\max(-1, 0)$

$\min(\frac{1}{3}, -\frac{2^{5/3}}{3})$

THE SECOND DERIVATIVE

The second derivative of a function can be used to determine where the function is concave upward and where it is concave downward. The second derivative can also be used to locate the points of inflection of a function, if there are any.

Concavity

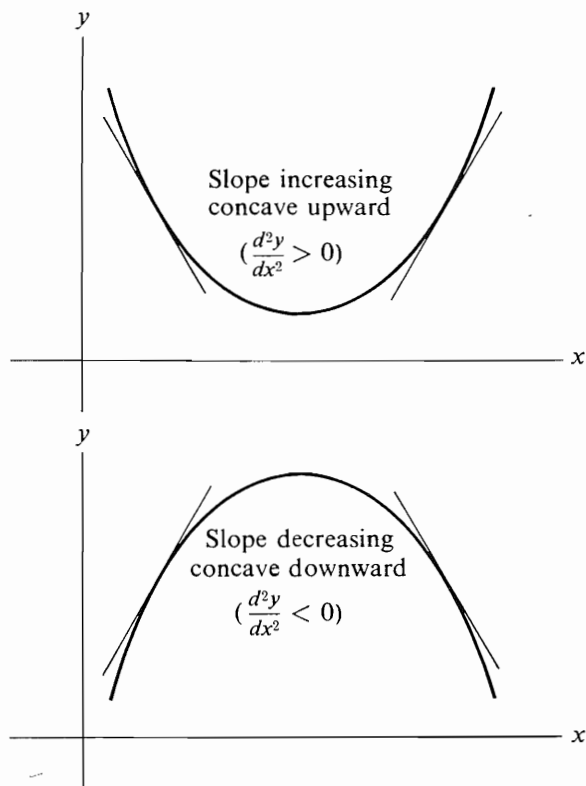
Consider the function $y = f(x)$ at the point $x = a$. The second derivative of y with respect to x , $y'' = f''(x)$, is the slope at the point x of the curve representing the first derivative $y' = f'(x)$ of the function $f(x)$; in particular, $f''(a)$ is the slope of the curve $y' = f'(x)$ at the point $x = a$.

If the second derivative $f''(a)$ of a function $y = f(x)$ is positive, $y' = f'(x)$ is an increasing function of x at $x = a$; the curve representing $y = f(x)$ is said to be *concave upward* where $f''(x)$ is positive. If the second derivative $f''(a)$ of a function $y = f(x)$ is negative, $y' = f'(x)$ is a decreasing function of x at $x = a$; the curve representing $y = f(x)$ is said to be *concave downward* where $f''(x)$ is negative (see Figure 2.42).

Consider a function $f(x)$ at a value $x = a$ for which $f(x)$ and $f'(x)$ are continuous. It is evident geometrically that if $f'(a) = 0$ and $f(x)$ is concave downward at $x = a$, then $f(x)$ has a maximum at a ; similarly, if $f'(a) = 0$ and $f(x)$ is concave upward at $x = a$, then $f(x)$ has a minimum at a . The corresponding algebraic argument is based on the fact that a positive second derivative indicates an increasing slope and a negative second derivative indicates a decreasing slope.

CAUTION: The second derivative test is only sufficient, not necessary, for a relative maximum or minimum, since $f(x)$ may be concave upward or downward at $x = a$ if $f''(a) = 0$. That is, if $f''(a) = 0$, then

Figure 2.42



$$f''(a) > 0 \Rightarrow \text{concave upward}$$

$$f''(a) < 0 \Rightarrow \text{concave downward}$$

From this follows a very convenient, although not always applicable, test for relative maximum and minimum values:

$$\text{If } \begin{cases} f(x) \text{ and } f'(x) \text{ are continuous} \\ \text{at } x = a \\ \text{and} \\ f'(a) = 0 \end{cases} \quad \text{then } \begin{cases} f''(a) > 0 \Rightarrow \text{relative minimum} \\ \text{at } x = a \\ f''(a) < 0 \Rightarrow \text{relative maximum} \\ \text{at } x = a \\ f''(a) = 0 \Rightarrow \text{test does not apply} \end{cases}$$

EXAMPLE

Find the relative maxima and minima (if any) of the function $y = \frac{1}{3}x^3 - 2x^2 + 3x + 1$.

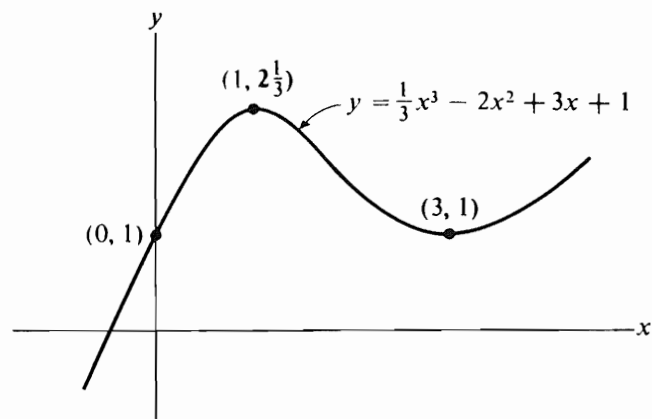
$$\frac{dy}{dx} = x^2 - 4x + 3$$

$$= 0 \quad \text{if } x^2 - 4x + 3 = 0$$

$$(x - 3)(x - 1) = 0$$

$$x = 3, 1$$

Figure 2.43



$$\frac{dy^2}{dx^2} = 2x - 4$$

If $x = 1$, $\frac{d^2y}{dx^2} < 0$, so maximum at $x = 1$

If $x = 3$, $\frac{d^2y}{dx^2} > 0$, so minimum at $x = 3$

See Figure 2.43.

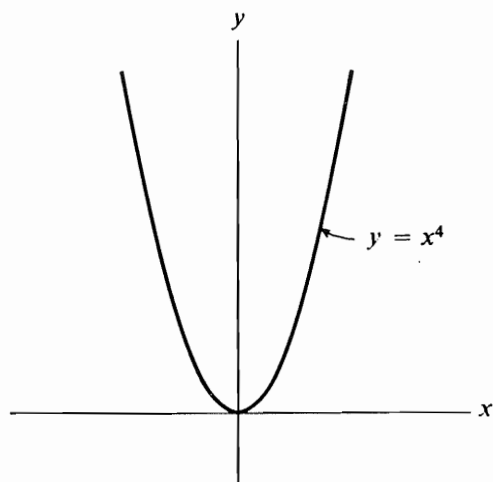
EXAMPLE

Find the relative maxima and minima (if any) of the function $y = x^4$.

$$\frac{dy}{dx} = 4x^3$$

$$= 0 \quad \text{if } x = 0$$

Figure 2.44



$$\frac{d^2y}{dx^2} = 12x^2$$

If $x = 0$, $\frac{d^2y}{dx^2} = 0$, so there may or may not be a maximum or minimum at $x = 0$

$$\left. \begin{array}{l} \text{If } x < 0, \frac{dy}{dx} < 0 \\ \text{If } x > 0, \frac{dy}{dx} > 0 \end{array} \right\} \text{so minimum at } x = 0$$

See Figure 2.44.

EXAMPLE

Find the relative maxima and minima (if any) of the function $y = x^3$.

$$\frac{dy}{dx} = 3x^2$$

$$= 0 \quad \text{if } x = 0$$

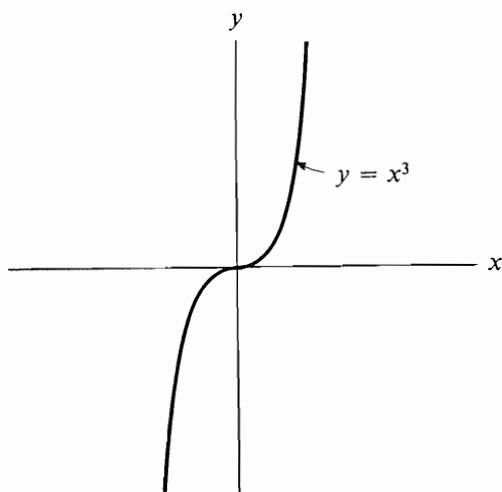
$$\frac{d^2y}{dx^2} = 6x$$

If $x = 0$, $\frac{d^2y}{dx^2} = 0$, so there may or may not be a maximum or minimum at $x = 0$

$$\left. \begin{array}{l} \text{If } x < 0, \frac{dy}{dx} > 0 \\ \text{If } x > 0, \frac{dy}{dx} > 0 \end{array} \right\} \text{so no maximum or minimum at } x = 0$$

See Figure 2.45.

Figure 2.45



In the examples above, when $f''(a) = 0$ —that is, when the second derivative test for a maximum or minimum does not apply—the first derivative is examined for change of sign as x passes through a . There is, however, an alternative procedure, involving derivatives of higher order, which is more convenient in some cases.

$$\text{If } \begin{cases} f'(a) = f''(a) = \cdots = f^{(n-1)}(a) = 0 \\ \text{and} \\ f^{(n)}(a) \neq 0 \end{cases} \quad \text{then } \begin{cases} \text{for } n \text{ even} & \begin{cases} f^{(n)}(a) < 0 \Rightarrow \text{relative} \\ & \text{maximum at } x = a \\ f^{(n)}(a) > 0 \Rightarrow \text{relative} \\ & \text{minimum at } x = a \end{cases} \\ \text{for } n \text{ odd, } f(a) \text{ is neither a} \\ & \text{relative maximum or minimum} \end{cases}$$

EXAMPLE

Find the relative maxima and minima (if any) of the function $y = 3x^4 - x^3 + 2$.

$$\begin{aligned} \frac{dy}{dx} &= 12x^3 - 3x^2 \\ &= 0 \quad \text{if } 12x^3 - 3x^2 = 0 \\ &\quad 3x^2(4x - 1) = 0 \\ &\quad x = 0, \frac{1}{4} \end{aligned}$$

$$\begin{aligned} \frac{d^2y}{dx^2} &= 36x^2 - 6x \\ &= 6x(6x - 1) \end{aligned}$$

$$\text{If } x = \frac{1}{4}, \frac{d^2y}{dx^2} > 0, \text{ so minimum at } x = \frac{1}{4}$$

$$\text{If } x = 0, \frac{d^2y}{dx^2} = 0, \text{ so there may or may not be a maximum or minimum at } x = 0$$

$$\begin{aligned} \frac{d^3y}{dx^3} &= 72x - 6 \\ &= 6(12x - 1) \end{aligned}$$

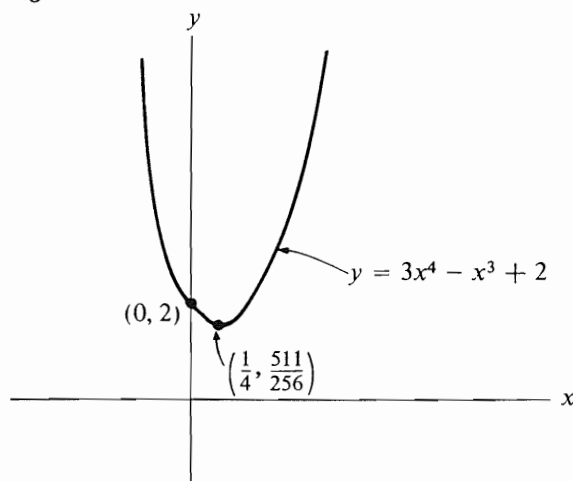
$$\text{If } x = 0, \frac{d^3y}{dx^3} \neq 0, \text{ so no maximum or minimum at } x = 0$$

Alternative procedure using change of sign of the first derivative:

$$\begin{aligned} &\left. \begin{array}{l} \text{If } x < 0, \frac{dy}{dx} < 0 \\ \text{If } 0 < x < \frac{1}{4}, \frac{dy}{dx} < 0 \end{array} \right\} \text{so no maximum or minimum at } x = 0 \\ &\left. \begin{array}{l} \text{If } 0 < x < \frac{1}{4}, \frac{dy}{dx} < 0 \\ \text{If } x > \frac{1}{4}, \frac{dy}{dx} > 0 \end{array} \right\} \text{so minimum at } x = \frac{1}{4} \end{aligned}$$

See Figure 2.46.

Figure 2.46



Points of Inflection

A function $y = f(x)$ is said to have a *point of inflection* at a point where the concavity of the curve changes. Geometrically, at a point of inflection the tangent crosses the curve. Since the sign of the second derivative indicates the concavity of a curve, it follows that a change in the sign of the second derivative implies a change in concavity (and thus a point of inflection).

If a function $f(x)$ has a point of inflection at a value $x = a$ for which its second derivative is continuous, then $f''(a) = 0$. Values of x for which $f''(x)$ is discontinuous must be considered separately, as discussed below.

CAUTIONS

1. A point of inflection at $x = a$ implies $f''(a) = 0$ only if $f(x)$ and $f''(x)$ are continuous at $x = a$.
2. $f''(a) = 0$ does not imply a point of inflection at $x = a$ even if $f(x)$ and $f''(x)$ are continuous at $x = a$; that is, if $f(x)$ and $f''(x)$ are continuous at $x = a$, $f''(a) = 0$ is a necessary but not a sufficient condition for a point of inflection at $x = a$.

Consider a function $f(x)$ at a value $x = a$ for which $f(x)$ and $f''(x)$ are continuous. If $f''(x)$ changes sign as x passes through the point $x = a$, then the curve representing $f(x)$ has a point of inflection at $x = a$. In this case $f''(a) = 0$.

A point of inflection may also occur at a value $x = a$ for which $f(x)$ is continuous but $f''(x)$ is discontinuous. Again, change in the sign of the second derivative as x passes through a is necessary for the existence of a point of inflection at $x = a$.

Thus, in order to determine the points of inflection (if any) of a function $y = f(x)$,

1. Determine the values of x for which $f''(x)$ is either zero or discontinuous and $f(x)$ is continuous.
2. For each value a such that $f(x)$ is continuous at $x = a$ and $f''(x)$ is either zero or discontinuous at $x = a$, determine whether $f''(x)$ changes sign as x increases through a :

$f''(x)$ changes sign at $x = a \Rightarrow$ point of inflection at $x = a$

$f''(x)$ does not change sign at $x = a \Rightarrow$ no point of inflection at $x = a$

NOTE: If $f(x)$ and $f''(x)$ are continuous at $x = a$ and $f''(a) = 0$, then $f'''(a) \neq 0$ implies a point of inflection at $x = a$; this may be used as an alternative to checking for change of sign of the second derivative at $x = a$.

EXAMPLE

Find the relative maxima and minima and the points of inflection (if any) of the function $y = x^{1/3}$.

$$\frac{dy}{dx} = \frac{1}{3}x^{-2/3}$$

$$\neq 0 \quad \text{for all } x$$

$$\frac{dy}{dx} \quad \text{discontinuous at } 0$$

$$\left. \begin{array}{l} \text{If } x < 0, \frac{dy}{dx} > 0 \\ \text{If } x > 0, \frac{dy}{dx} > 0 \end{array} \right\} \text{so no maximum or minimum at } x = 0$$

$$\frac{d^2y}{dx^2} = -\frac{2}{9}x^{-5/3}$$

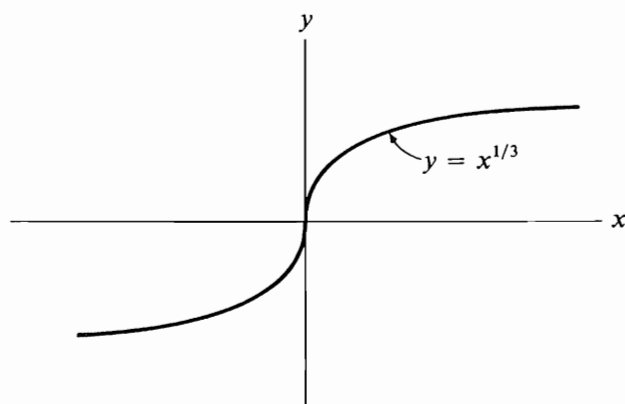
$$\neq 0 \quad \text{for all } x$$

$$\frac{d^2y}{dx^2} \quad \text{discontinuous at } 0$$

$$\left. \begin{array}{l} \text{If } x < 0, \frac{d^2y}{dx^2} > 0 \\ \text{If } x > 0, \frac{d^2y}{dx^2} < 0 \end{array} \right\} \text{so point of inflection at } x = 0$$

See Figure 2.47.

Figure 2.47



EXAMPLE

Find the relative maxima and minima and the points of inflection (if any) of the function $y = x/\sqrt{1-x^2}$.

$$\frac{dy}{dx} = -\frac{1}{2}x(1-x^2)^{-3/2}(-2x) + (1-x^2)^{-1/2}$$

$$= (1-x^2)^{-3/2}(x^2 + 1 - x^2)$$

$$= (1-x^2)^{-3/2}$$

$$\frac{dy}{dx} \rightarrow \infty \text{ as } x \rightarrow \pm 1, \text{ so infinite discontinuities at } x = \pm 1$$

$$\frac{dy}{dx} \geq 0 \text{ for all } x, -1 \leq x \leq 1, \text{ so no maximum or minimum}$$

(Note that $y = x/\sqrt{1-x^2}$ is defined only for $-1 \leq x \leq 1$.)

$$\frac{d^2y}{dx^2} = -\frac{3}{2}(1-x^2)^{-5/2}(-2x)$$

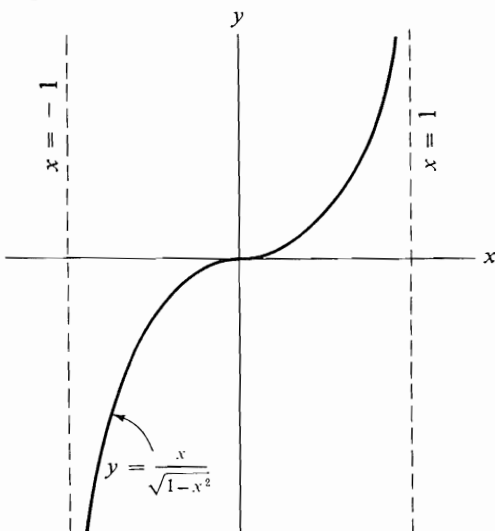
$$= 3x(1-x^2)^{-5/2}$$

$$= 0 \quad \text{if} \quad x = 0$$

$$\left. \begin{array}{l} \text{If } -1 < x < 0, \frac{d^2y}{dx^2} < 0 \\ \text{If } 0 < x < 1, \frac{d^2y}{dx^2} > 0 \end{array} \right\} \text{so point of inflection at } x = 0$$

See Figure 2.48.

Figure 2.48

**EXAMPLE**

Find the relative maxima and minima and the points of inflection (if any) of the function $y = x + (1/x)$.

$$\frac{dy}{dx} = 1 - \frac{1}{x^2}$$

$$= \frac{x^2 - 1}{x^2}$$

$$= 0 \quad \text{if } x = \pm 1$$

$\frac{dy}{dx}$ discontinuous at $x = 0$

If $x < -1$, $\frac{dy}{dx} > 0$ }
 If $-1 < x < 0$, $\frac{dy}{dx} < 0$ } so maximum at $x = -1$

If $0 < x < 1$, $\frac{dy}{dx} < 0$ }
 If $x > 1$, $\frac{dy}{dx} > 0$ } so minimum at $x = 1$

If $-1 < x < 0$, $\frac{dy}{dx} < 0$ }
 If $0 < x < 1$, $\frac{dy}{dx} < 0$ } so no maximum or minimum at $x = 0$

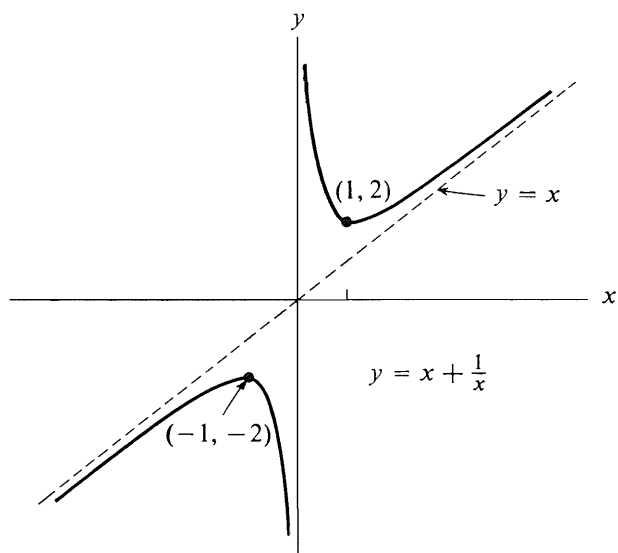
$$\frac{d^2y}{dx^2} = \frac{2}{x^3}$$

$$\neq 0 \quad \text{for all } x$$

$\frac{d^2y}{dx^2}$ discontinuous at $x = 0$ but $\frac{dy}{dx}$ also discontinuous at $x = 0$, so no point of inflection at $x = 0$

See Figure 2.49.

Figure 2.49



SUMMARY OF THE INFORMATION THAT CAN BE OBTAINED FROM THE SECOND DERIVATIVE

Consider the function $y = f(x)$ at the point $x = a$:

$$f''(a) > 0 \Rightarrow f(x) \text{ concave upward at } x = a$$

$$f''(a) < 0 \Rightarrow f(x) \text{ concave downward at } x = a$$

$$\begin{array}{ll} \text{If } \left\{ \begin{array}{l} f(x) \text{ and } f'(x) \text{ are continuous} \\ \text{at } x = a \\ \text{and} \\ f'(a) = 0 \end{array} \right. & \text{then } \left\{ \begin{array}{l} f''(a) > 0 \Rightarrow \text{relative minimum at } x = a \\ f''(a) < 0 \Rightarrow \text{relative maximum at } x = a \\ f''(a) = 0 \Rightarrow \text{test does not apply} \end{array} \right. \\ \\ \text{If } \left\{ \begin{array}{l} f(x) \text{ is continuous at } x = a \\ \text{and either} \\ f''(a) = 0 \\ \text{or} \\ f''(x) \text{ is discontinuous at } x = a \end{array} \right. & \text{then } \left\{ \begin{array}{l} f''(x) \text{ changes sign at } x = a \\ \Rightarrow \text{point of inflection at } x = a \\ f''(x) \text{ does not change sign at } x = a \\ \Rightarrow \text{no point of inflection at } x = a \end{array} \right. \end{array}$$

SUMMARY OF THE PROCEDURE FOR SKETCHING THE GRAPH OF $y = f(x)$ USING INFORMATION PROVIDED BY FIRST AND SECOND DERIVATIVES

1. Obtain $\frac{dy}{dx}$ and $\frac{d^2y}{dx^2}$.
2. Determine the ranges of values of x for which $\frac{dy}{dx}$ is positive and for which it is negative.
3. Check points where $\frac{dy}{dx} = 0$ for possible maxima or minima.
4. Check points where $\frac{dy}{dx}$ is discontinuous for possible maxima or minima.
5. Determine the ranges of values of x for which $\frac{d^2y}{dx^2}$ is positive and for which it is negative.
6. Check points where $\frac{d^2y}{dx^2} = 0$ for possible points of inflection.
7. Check points where $\frac{d^2y}{dx^2}$ is discontinuous for possible points of inflection.
8. Determine the intercepts.
9. Note the nature of the curve for very small and very large values of x .
10. Sketch the curve as indicated by the signs of $\frac{dy}{dx}$ (increasing or decreasing) and of $\frac{d^2y}{dx^2}$ (concave upward or downward), the intercepts, asymptotic properties, and discontinuities.

EXAMPLE

Sketch the curve represented by the function $y = 4 + 3x - x^3$.

$$\frac{dy}{dx} = 3 - 3x^2$$

$$\frac{d^2y}{dx^2} = -6x$$

$$\frac{dy}{dx} = 0 \quad \text{if } x = \pm 1$$

$$\text{If } x < -1, \frac{dy}{dx} < 0 \quad (\text{decreasing})$$

$$\text{If } -1 < x < 1, \frac{dy}{dx} > 0 \quad (\text{increasing})$$

$$\text{If } x > 1, \frac{dy}{dx} < 0 \quad (\text{decreasing})$$

so minimum at $(-1, 2)$ and maximum at $(1, 6)$. (As a check, if $x = -1$, $\frac{d^2y}{dx^2} = 6$, and if $x = 1$, $\frac{d^2y}{dx^2} = -6$.)

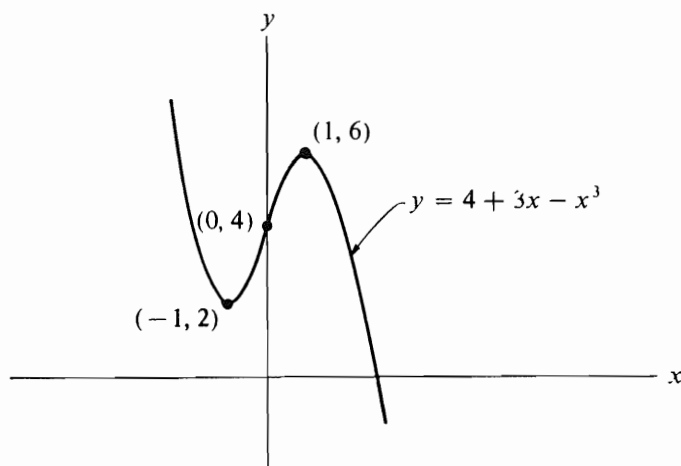
$$\frac{d^2y}{dx^2} = 0 \quad \text{if } x = 0$$

$$> 0 \quad \text{if } x < 0 \quad (\text{concave upward})$$

$$< 0 \quad \text{if } x > 0 \quad (\text{concave downward})$$

so point of inflection at $(0, 4)$. If $x = 0$, $y = 4$. As $x \rightarrow \infty$, $y \rightarrow -\infty$, and as $x \rightarrow -\infty$, $y \rightarrow \infty$. See Figure 2.50.

Figure 2.50



EXAMPLE

Sketch the curve represented by the function $y = x + (4/x)$.

$$\frac{dy}{dx} = 1 - \frac{4}{x^2}$$

$$\frac{d^2y}{dx^2} = \frac{8x}{x^4} - \frac{3}{x^3}$$

$$\frac{dy}{dx} = 0 \quad \text{if } x = \pm 2$$

$$\text{If } x < -2, \frac{dy}{dx} > 0 \quad (\text{increasing})$$

$$\text{If } -2 < x < 2 \text{ and } x \neq 0, \frac{dy}{dx} < 0 \quad (\text{decreasing})$$

$$\text{If } x > 2, \frac{dy}{dx} > 0 \quad (\text{increasing})$$

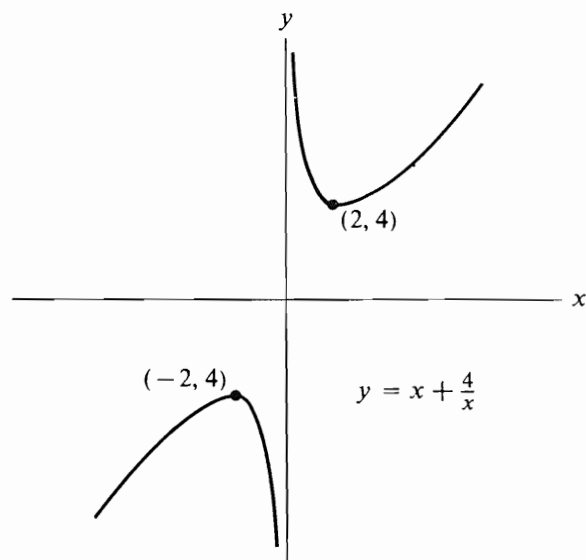
so minimum at $(2, 4)$ and maximum at $(-2, -4)$. y , $\frac{dy}{dx}$, and $\frac{d^2y}{dx^2}$ discontinuous at $x = 0$ (if $x = 2$, $\frac{d^2y}{dx^2} = 1$, and if $x = -2$, $\frac{d^2y}{dx^2} = -1$).

$$-\frac{d^2y}{dx^2} < 0 \quad \text{if } x < 0 \quad (\text{concave downward})$$

$$> 0 \quad \text{if } x > 0 \quad (\text{concave upward})$$

No point of inflection at $x = 0$ because of discontinuity of $y = x + (4/x)$. No x -intercepts or y -intercepts. As $x \rightarrow \infty$, $y \rightarrow \infty$, and as $x \rightarrow -\infty$, $y \rightarrow -\infty$. See Figure 2.51.

Figure 2.51



EXAMPLE

Find the points at which the curve represented by the function $x - y^4 + 2y^2 - 2 = 0$ has vertical tangents and sketch the curve.

$$1 - 4y^3 \frac{dy}{dx} + 4y \frac{dy}{dx} = 0$$

$$\frac{dy}{dx} = \frac{1}{4y(y^2 - 1)}$$

so $\frac{dy}{dx} \neq 0$ for all x . Vertical tangents occur when $\frac{dy}{dx}$ is undefined, so there are vertical tangents at $(2, 0)$, $(1, -1)$, and $(1, 1)$. (Note that $x = y^4 - 2y^2 + 2$.)

$$\frac{d^2y}{dx^2} = \frac{-(12y^2 - 4) \frac{dy}{dx}}{[4y(y^2 - 1)]^2} = \frac{-4(3y^2 - 1)}{[4y(y^2 - 1)]^3}$$

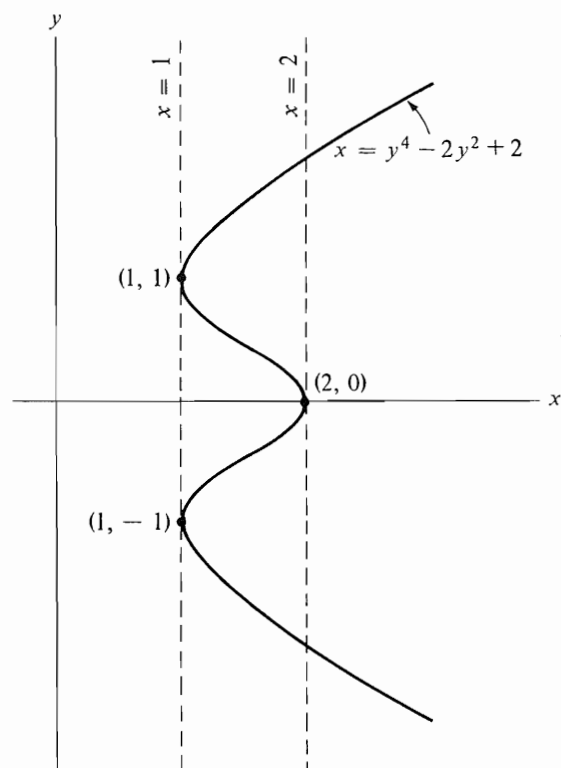
$$\frac{d^2y}{dx^2} = 0 \quad \text{if } y = \pm\sqrt{\frac{1}{3}}$$

$$\text{If } y < -1, \frac{d^2y}{dx^2} > 0 \quad (\text{concave upward})$$

$$\text{If } -1 < y < -\sqrt{\frac{1}{3}}, \frac{d^2y}{dx^2} < 0 \quad (\text{concave downward})$$

$$\text{If } -\sqrt{\frac{1}{3}} < y < 0, \frac{d^2y}{dx^2} > 0 \quad (\text{concave upward})$$

Figure 2.52



so there is an inflection point at $y = -\sqrt{\frac{1}{3}}$.

If $0 < y < \sqrt{\frac{1}{3}}$, $\frac{d^2y}{dx^2} < 0$ (concave downward)

If $\sqrt{\frac{1}{3}} < y < 1$, $\frac{d^2y}{dx^2} > 0$ (concave upward)

so there is an inflection point at $y = \sqrt{\frac{1}{3}}$.

If $y > 1$, $\frac{d^2y}{dx^2} < 0$ (concave downward)

If $y = 0$, $x = 2$. As $y \rightarrow \infty$, $x \rightarrow \infty$, and as $y \rightarrow -\infty$, $x \rightarrow \infty$. See Figure 2.52.

PROBLEMS

For each of the following functions, determine any maxima, minima, and points of inflection; sketch the curve representing each function. (Refer to problems on pages 230–231 for maxima and minima.)

1. $y = 12 - 12x + x^3$
2. $y = \frac{x}{x+1}$
3. $y = \frac{x^3}{3} - \frac{x^2}{2} - 6x$
4. $y = x^4 - 32x + 48$
5. $y = \frac{x}{\sqrt{x^2+7}}$
6. $y = \frac{2x}{\sqrt{x^2-1}}$
7. $y = x^2 - 4x + 3$
8. $y = x\sqrt{1-x^2}$
9. $y = x^3 - 3x^2 + 2$
10. $y = \frac{1}{x^2+4}$
11. $y = \frac{x^2}{\sqrt{x^2-8}}$
12. $y = x^4 - 4x^3 + 12$
13. $y = \frac{3x^2}{\sqrt{x^2+3}}$
14. $y = \frac{1}{16-x^2}$
15. $y = \frac{2}{3}x^3 - 4x^2 + 6x + 2$
16. $y = \frac{8x}{x^2+4}$
17. $y = \frac{x+3}{x^2}$
18. $y = x^5 + 6$
19. $y = (x-1)^{1/3}(x+1)^{2/3}$
20. $y = \frac{1}{6}(x^3 - 6x^2 + 9x + 6)$
21. Sketch a smooth curve $y = f(x)$ having all the following properties:
 - a. $f(1) = 0$
 $f'(x) < 0$ for $x < 1$
 $f'(x) > 0$ for $x > 1$
 - b. $f(1) = 0$
 $f''(x) < 0$ for $x < 1$
 $f''(x) > 0$ for $x > 1$
22. Sketch a smooth curve $y = f(x)$ having all the following properties:

$f(0) = 10$	$f'(6) = 0$	$f''(x) < 0$ for $x < 9$
$f(6) = 15$	$f'(10) = 0$	$f''(9) = 0$
$f(10) = 0$		$f''(x) > 0$ for $x > 9$
23. Sketch a continuous curve $y = f(x)$ having all of the following properties:

$f(-2) = 8$	$f'(x) > 0$ for $ x > 2$	$f''(x) < 0$ for $x < 0$
$f(0) = 4$	$f'(2) = f'(-2) = 0$	$f''(x) > 0$ for $x > 0$
$f(2) = 0$		
24. Sketch a continuous curve $y = f(x)$ having the properties $f''(x) > 0$ for $x < 2$ and $f''(x) < 0$ for $x > 2$:
 - a. If $f'(x)$ is continuous at $x = 2$
 - b. If $f'(x) \rightarrow 1$ as $x \rightarrow 2^-$ and $f'(x) \rightarrow -1$ as $x \rightarrow 2^+$
 - c. If $f'(x) = 1$ for $x < 2$ and $f'(x) = -1$ for $x > 2$

25. Sketch a continuous curve $y = f(x)$ having the following properties:
- | | |
|--------------------------------------|--------------------------|
| a. $f(2) = f'(2) = 0$ | b. $f(2) = f'(2) = 0$ |
| $f(0) = 2$ | $f(0) = 2$ |
| $f''(x) < 0$ for $x < 1$ and $x > 3$ | $f''(x) > 0$ for all x |
| $f''(x) > 0$ for $1 < x < 3$ | |
26. Sketch a continuous curve $y = f(x)$ having all of the following properties:
- | | | |
|-------------|-------------------------|--------------------------|
| $f(0) = 10$ | $f'(x) > 0$ for $x < 0$ | $f''(x) > 0$ for $x < 0$ |
| $f(-3) = 0$ | $f'(x) < 0$ for $x > 0$ | $f''(x) > 0$ for $x > 0$ |
| $f(3) = 0$ | | |
27. Show that a cubic curve whose equation is of the form
- $$y = ax^3 + bx^2 + cx + d \quad a, b, c, d \neq 0$$
- has one (and only one) point of inflection.
28. Sketch a continuous curve $y = f(x)$ for $x > 0$ if $f(1) = 0$ and $f'(x) = \frac{1}{x}$ for $x > 0$. Is such a curve necessarily concave upward or downward?

The curves representing the functions discussed on pages 221–229 were sketched without discussion of points of inflection; investigate points of inflection and sketch the curves accurately.

Answers to Odd-Numbered Problems

- | | |
|--|---|
| 1. max $(-2, 28)$ | 13. no max |
| min $(2, -4)$ | min $(0, 0)$ |
| pt. inf. $(0, 12)$ | pt. inf. $(\pm\sqrt{6}, 6)$ |
| 3. max $(-2, \frac{22}{3})$ | 15. max $(1, \frac{14}{3})$ |
| min $(3, -\frac{27}{2})$ | min $(3, 2)$ |
| pt. inf. $(\frac{1}{2}, -\frac{37}{12})$ | pt. inf. $(2, \frac{10}{3})$ |
| 5. no max | 17. no max |
| no min | min $(-6, -\frac{1}{2})$ |
| pt. inf. $(0, 0)$ | pt. inf. $(-9, -\frac{2}{27})$ |
| 7. no max | 19. max $(-1, 0)$ |
| min $(2, -1)$ | min $(\frac{1}{3}, -\frac{2^{5/3}}{3})$ |
| no pt. inf. | pt. inf. $(1, 0)$ |
| 9. max $(0, 2)$ | 27. pt. inf. at $x = -\frac{b}{3a}$ |
| min $(2, -2)$ | |
| pt. inf. $(1, 0)$ | |
| 11. no max | |
| min $(\pm 4, 4\sqrt{2})$ | |
| no pt. inf. | |

2.12 Applications of Derivatives in Business and Economics

In this section several applications of derivatives in business and economics are discussed. Marginal cost, elasticity and marginal revenue are considered in much more detail than in previous sections. In addition, revenue from taxation, profit under monopoly, the effect of taxation on monopoly and inventory models are discussed.

COST, AVERAGE COST, AND MARGINAL COST

If the total cost y of producing and marketing x units of a commodity is assumed to be a function of x only, then the total cost function can be represented by $y = f(x)$.

Functions of various types are used to represent total cost curves; in general, cost curves have the following properties:

1. When no units are produced, total cost is zero or positive—that is, $f(0) \geq 0$. If $f(0) > 0$, then $f(0)$ is amount of overhead or fixed costs of production.
2. Total cost increases as x increases, so $f'(x)$ is always positive.
3. The cost of producing an extremely large quantity of any commodity usually reaches a point at which it increases at an increasing rate. Thus usually the total cost curve eventually is concave upward, that is, $f''(x) > 0$; however, in a limited range the total cost curve is frequently concave downward, corresponding to decreasing marginal cost.

If the total cost function is

$$y = f(x)$$

then the average or per unit cost is

$$\bar{y} = \frac{y}{x} = \frac{f(x)}{x}$$

and the marginal cost is

$$\frac{dy}{dx} = f'(x)$$

The first derivative of average cost (the marginal average cost) is

$$\frac{d\bar{y}}{dx} = \frac{xf'(x) - f(x)}{x^2}$$

$$= 0 \quad \text{if and only if } xf'(x) - f(x) = 0$$

$$\text{that is, } f(x) = xf'(x),$$

$$\frac{f(x)}{x} = f'(x)$$

Thus average cost is minimum at the value of x for which average cost equals marginal cost; that is, the average and marginal cost curves intersect at the point of minimum average cost. Note that the value of x (if any) for which $\frac{d\bar{y}}{dx} = 0$ is assumed to be a minimum because of the third property of cost curves mentioned above; for a particular total cost curve the existence of this minimum can be checked in the usual way.

EXAMPLE

Linear cost functions:

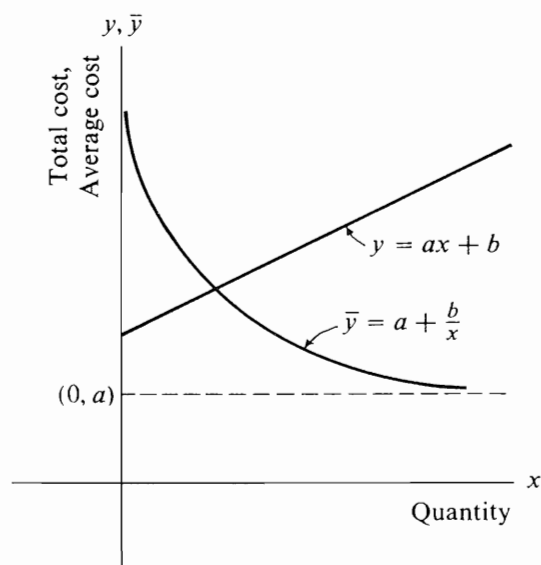
$$\text{Total cost: } y = ax + b, \text{ where } a > 0, b \geq 0$$

$$\text{Average cost: } \bar{y} = y/x = a + (b/x)$$

$$\text{Marginal cost: } \frac{dy}{dx} = a$$

$$\text{Marginal average cost: } \frac{d\bar{y}}{dx} = -\frac{b}{x^2}$$

Figure 2.53



Total cost and marginal cost are represented by straight lines and average cost is represented by the first quadrant branch of an equilateral hyperbola with horizontal asymptote $\bar{y} = a$. Average cost is thus a decreasing function of the number of units produced; it has no minimum value but approaches a , the marginal cost, as the number of units increases (see Figure 2.53).

EXAMPLE

Quadratic cost functions:

Total cost: $y = ax^2 + bx + c$, where $a > 0$, $b \geq 0$, $c \geq 0$

Average cost: $\bar{y} = y/x = ax + b + (c/x)$

Marginal cost: $\frac{dy}{dx} = 2ax + b$

Marginal average cost: $\frac{d\bar{y}}{dx} = a - \frac{c}{x^2}$

Total cost is represented by the first quadrant part of a parabola, average cost is represented by the first quadrant branch of a hyperbola, and marginal cost is represented by a straight line (see Figure 2.54).

$$\frac{d\bar{y}}{dx} = 0 \quad \text{if } ax^2 = c$$

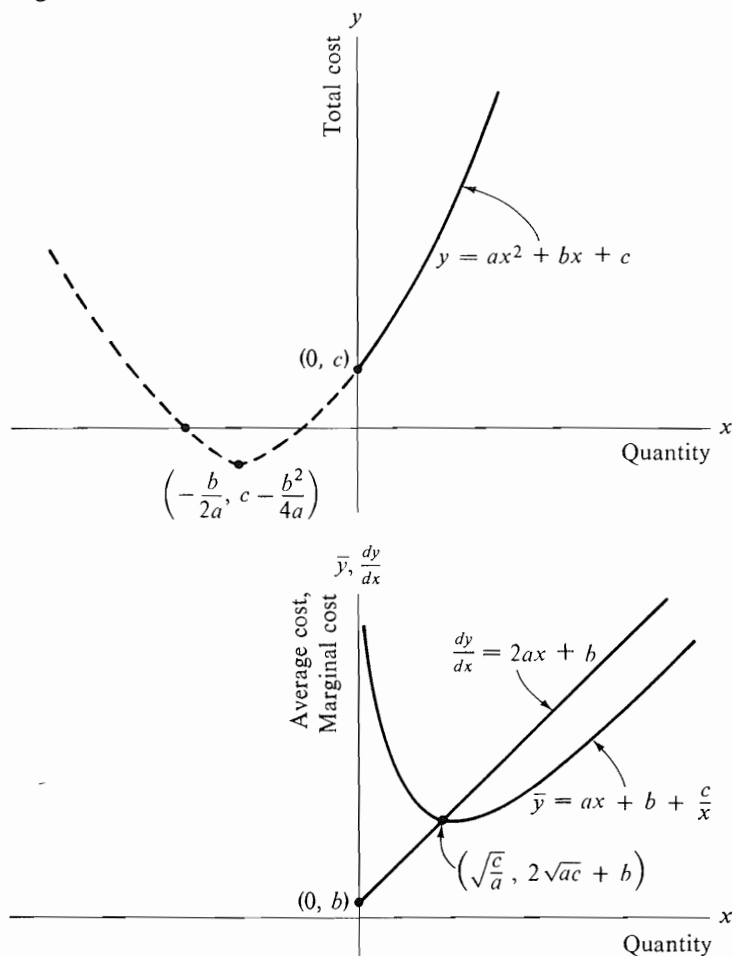
$$x = \pm \sqrt{\frac{c}{a}}, \text{ but only } x = \sqrt{\frac{c}{a}} \text{ is of interest}$$

$$\frac{d^2\bar{y}}{dx^2} = -\frac{-2cx}{x^4} = \frac{2c}{x^3} > 0 \text{ for } x > 0, \text{ so minimum at } x = \sqrt{\frac{c}{a}}$$

Note that average and marginal cost are equal if $x = \sqrt{c/a}$, since

$$\begin{aligned} \bar{y} &= a\sqrt{\frac{c}{a}} + b + c\sqrt{\frac{a}{c}} \\ &= 2\sqrt{ac} + b \end{aligned}$$

Figure 2.54



and

$$\begin{aligned}\frac{dy}{dx} &= 2a\sqrt{\frac{c}{a}} + b \\ &= 2\sqrt{ac} + b\end{aligned}$$

EXAMPLE

Cubic cost functions:

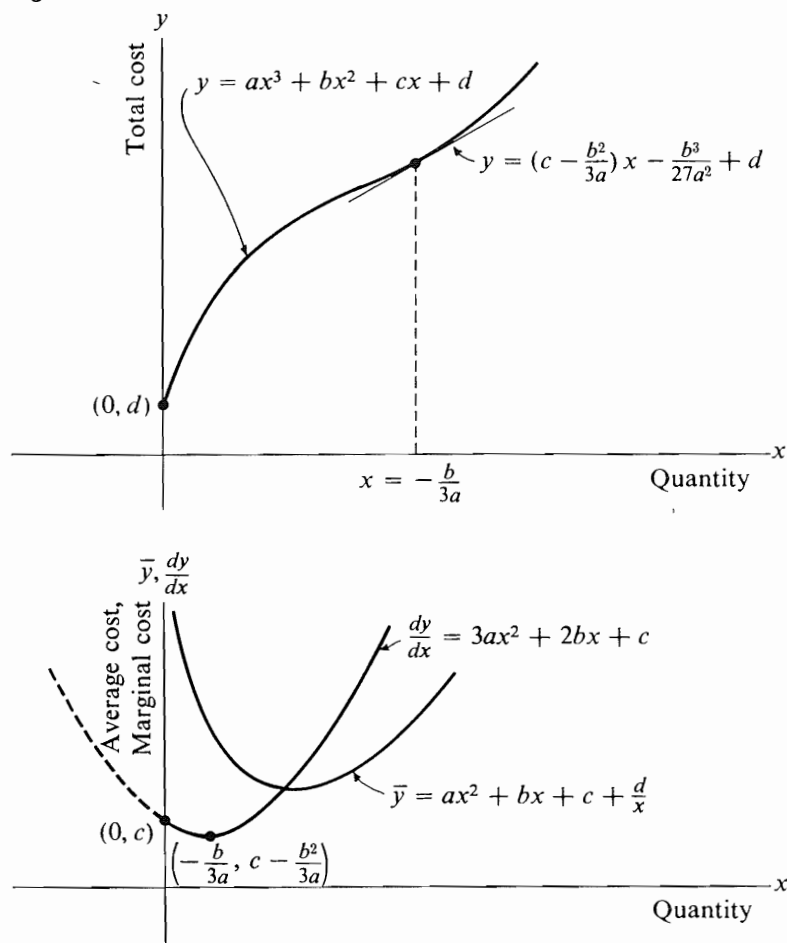
Total cost: $y = ax^3 + bx^2 + cx + d$, where $a > 0$, $b \leq 0$, $c \geq 0$, $d \geq 0$, $b^2 \leq 3ac$

Average cost: $\bar{y} = \frac{y}{x} = ax^2 + bx + c + \frac{d}{x}$

Marginal cost: $\frac{dy}{dx} = 3ax^2 + 2bx + c$

Marginal average cost: $\frac{d\bar{y}}{dx} = 2ax + b - \frac{d}{x^2}$

Figure 2.55



Total cost is represented by the first quadrant part of a cubic curve, average cost is represented by the first quadrant branch of a hyperbola, and marginal cost is represented by the first quadrant part of a parabola (see Figure 2.55).

Cubic cost functions are frequently appropriate when the concavity of the cost function changes in a specified interval; however, such a function must have no relative maximum or minimum in the first quadrant:

$$\frac{dy}{dx} = 3ax^2 + 2bx + c$$

$$= 0$$

$$\text{if } 3ax^2 + 2bx + c = 0$$

$$x = \frac{-2b \pm \sqrt{4b^2 - 12ac}}{6a} = \frac{-b \pm \sqrt{b^2 - 3ac}}{3a}$$

If $b^2 - 3ac < 0$, $y = ax^3 + bx^2 + cx + d$ has no relative maximum or minimum.

If $b^2 - 3ac = 0$, $x = -b/3a$ and

$$\begin{aligned} y &= a\left(\frac{-b^3}{27a^3}\right) + b\left(\frac{b^2}{9a^2}\right) + c\left(\frac{-b}{3a}\right) + d \\ &= \frac{2b^3}{27a^2} - \frac{bc}{3a} + d \\ &= \frac{b(2b^2 - 9ac)}{27a^2} + d \end{aligned}$$

$2b^2 - 9ac < 0$, since $b^2 - 3ac < 0$ and thus, since $a > 0$ and $b \leq 0$, $b(2b^2 - 9ac)/27a^2 \geq 0$ and $\{-b/3a, [b(2b^2 - 9ac)/27a^2] + d\}$ lies in the first quadrant.

$$\frac{d^2y}{dx^2} = 6ax + 2b$$

$$\left. \frac{d^2y}{dx^2} \right|_{x=-b/3a} = 0$$

$$\frac{d^3y}{dx^3} = 6a \neq 0$$

So no relative maximum or minimum at $x = -b/3a$ but a point of inflection at $x = -b/3a$. Thus, if $b^2 - 3ac \leq 0$, then $y = ax^3 + bx^2 + cx + d$ has no relative maximum or minimum in the first quadrant; a point of inflection occurs at

$$x = -\frac{b}{3a}, \quad y = \frac{b(2b^2 - 9ac)}{27a^2} + d$$

Note that $b^2 - 3ac \leq 0$ is a sufficient but not a necessary condition for $y = ax^3 + bx^2 + cx + d$ to have no relative maximum or minimum in the first quadrant. Other cases are discussed in Technical Note II.

Frequently the tangent at the point of inflection provides a good linear approximation of a cubic cost function for a limited range. The tangent at the point of inflection $x = -b/3a$ has slope

$$\frac{dy}{dx} = 3ax^2 + 2bx + c$$

$$\left. \frac{dy}{dx} \right|_{x=-b/3a} = \frac{b^2}{3a} - \frac{2b^2}{3a} + c$$

$$= c - \frac{b^2}{3a} \quad \text{which is } \geq 0, \text{ since } b^2 \leq 3ac$$

If $x = -b/3a$,

$$\begin{aligned} y &= a\left(-\frac{b^3}{27a^3}\right) + b\left(\frac{b^2}{9a^2}\right) + c\left(-\frac{b}{3a}\right) + d \\ &= \frac{2b^3}{27a^2} - \frac{bc}{3a} + d \end{aligned}$$

The line through the point $(-b/3a, 2b^3/27a^2 - (bc/3a) + d)$ and having slope $c - (b^2/3a)$ is given by

$$\begin{aligned} y - y_1 &= m(x - x_1) \\ y - \left(\frac{2b^3}{27a^2} - \frac{bc}{3a} + d\right) &= \left(c - \frac{b^2}{3a}\right)\left(x + \frac{b}{3a}\right) \\ y &= \left(c - \frac{b^2}{3a}\right)x + \frac{bc}{3a} - \frac{b^3}{9a^2} + \frac{2b^3}{27a^2} - \frac{bc}{3a} + d \\ &= \left(c - \frac{b^2}{3a}\right)x - \frac{b^3}{27a^2} + d \end{aligned}$$

EXAMPLE

Higher-order polynomial cost functions:

Total cost: $y = ax^b + c$, where $a > 0$, $b > 1$, $c \geq 0$

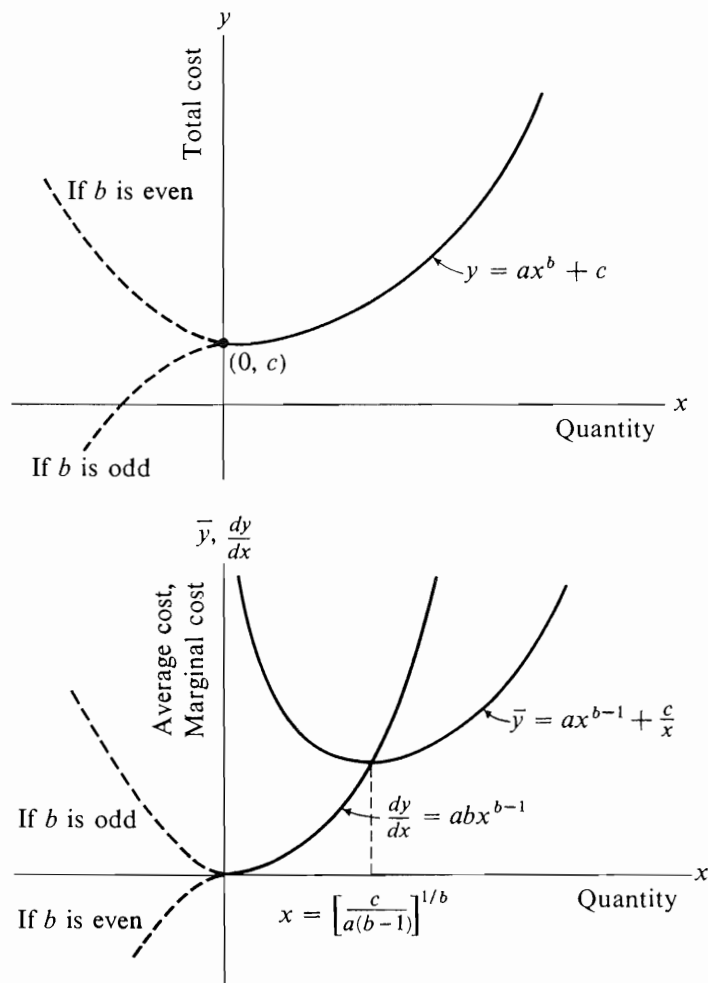
Average cost: $\bar{y} = \frac{y}{x} = ax^{b-1} + \frac{c}{x}$

Marginal cost: $\frac{dy}{dx} = abx^{b-1}$

Marginal average cost: $\frac{d\bar{y}}{dx} = a(b-1)x^{b-2} - \frac{c}{x^2}$

Total cost is represented by the first quadrant part of an algebraic curve, average cost is represented by the first quadrant branch of a hyperbola, and marginal cost is represented by the first quadrant part of an algebraic curve (see Figure 2.56).

Figure 2.56



This type of simple polynomial curve is always concave upward for $x \geq 0$ and thus in practice is usually appropriate only for limited ranges of x .

$$\frac{d\bar{y}}{dx} = 0 \quad \text{if } a(b-1)x^b = c$$

$$x = \left[\frac{c}{a(b-1)} \right]^{1/b}$$

$$\begin{aligned} \frac{d\bar{y}^2}{dx^2} &= a(b-1)(b-2)x^{b-3} + \frac{2c}{x^3} \\ &= x^{-3}[a(b-1)(b-2)x^b + 2c] \end{aligned}$$

If $x = \left[\frac{c}{a(b-1)} \right]^{1/b}$, then

$$\begin{aligned} \frac{d\bar{y}^2}{dx^2} &= \left[\frac{c}{a(b-1)} \right]^{-3/b} \left[a(b-1)(b-2) \frac{c}{a(b-1)} + 2c \right] \\ &> 0 \end{aligned}$$

so minimum at $x = \left[\frac{c}{a(b-1)} \right]^{1/b}$.

Note that average cost and marginal cost are equal if $x = (c/a(b-1))^{1/b}$, since

$$\begin{aligned} \bar{y} &= a \left(\frac{c}{a(b-1)} \right)^{(b-1)/b} + c \left(\frac{c}{a(b-1)} \right)^{-1/b} \\ &= \left(\frac{c}{a(b-1)} \right)^{(b-1)/b} \left[a + c \left(\frac{c}{a(b-1)} \right)^{-1} \right] \\ &= \left(\frac{c}{a(b-1)} \right)^{(b-1)/b} [a + a(b-1)] \\ &= ab \left(\frac{c}{a(b-1)} \right)^{(b-1)/b} \end{aligned}$$

and

$$\frac{dy}{dx} = ab \left(\frac{c}{a(b-1)} \right)^{(b-1)/b}$$

EXAMPLE

Exponential cost functions:

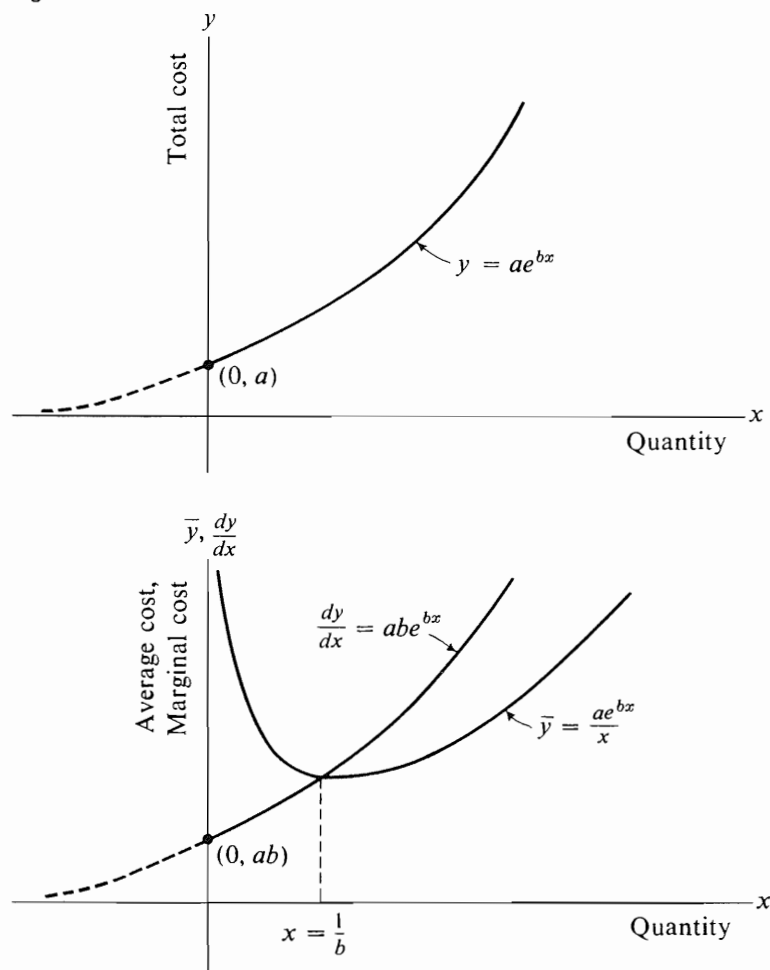
Total cost: $y = ae^{bx}$, where $a > 0$, $b > 0$

Average cost: $\bar{y} = \frac{y}{x} = \frac{ae^{bx}}{x}$

Marginal cost: $\frac{dy}{dx} = abe^{bx}$

Marginal average cost: $\frac{d\bar{y}}{dx} = \frac{abxe^{bx} - ae^{bx}}{x^2} = \frac{ae^{bx}(bx-1)}{x^2}$

Figure 2.57



Total cost is represented by the first quadrant part of an exponential curve, average cost is represented by the first quadrant branch of an exponential hyperbola, and marginal cost is represented by the first quadrant part of an exponential curve (see Figure 2.57).

Exponential and logarithmic curves are either concave upward or concave downward for all $x \geq 0$ and thus in practice are usually appropriate only for limited ranges of x .

$$\frac{d\bar{y}}{dx} = 0 \quad \text{if } ae^{bx}(bx - 1) = 0$$

$$x = \frac{1}{b}$$

$$\begin{aligned} \frac{d^2\bar{y}}{dx^2} &= \frac{(abe^{bx} + ab^2xe^{bx} - abe^{bx})x^2 - (abxe^{bx} - ae^{bx})(2x)}{x^4} \\ &= \frac{axe^{bx}(b^2x^2 - 2bx + 2)}{x^4} \end{aligned}$$

If $x = 1/b$, then

$$\begin{aligned} \frac{d^2\bar{y}}{dx^2} &= \frac{\frac{a}{b}e(1 - 2 + 2)}{\frac{1}{b^4}} \\ &= ab^3e > 0 \end{aligned}$$

so minimum at $x = 1/b$. Note that average cost and marginal cost are equal if $x = 1/b$, since

$$\bar{y} = abe$$

and

$$\frac{dy}{dx} = abe$$

PROBLEMS

1. For each of the following average cost functions, find the minimum average cost and show that at the minimum average cost, marginal cost and average cost are equal.

a. $\bar{y} = 25 - 8x + x^2$	e. $\bar{y} = 2x + 5 + \frac{18}{x}$
b. $\bar{y} = 2 + x \ln x$	f. $\bar{y} = 20 + 2x^2 + 4x^4$
c. $\bar{y} = 2e^x + e^{-x}$	g. $\bar{y} = 10 - 4x^3 + 3x^4$
d. $\bar{y} = 3x + 5 + \frac{6}{x}$	h. $\bar{y} = 6x + 7 + \frac{36}{x}$

2. For each of the following total cost functions, find marginal cost and determine the nature of marginal cost (increasing, decreasing).

a. $y = 1000x - 180x^2 + 3x^3$ b. $y = 220 + 55x - 2x^3 + x^4$

3. Determine the nature of the marginal and average cost functions (increasing, decreasing) for each of the following total cost functions.

a. $y = \sqrt{x + 25}$, $0 \leq x \leq 10$ b. $y = 9x + 5xe^{-2x}$

4. For the following total cost function, find the equation of the tangent at the point of inflection as an approximation of the function near that point.

$$y = x^3 - 6x^2 + 14x + 6$$

5. The Precision Machine Tool Manufacturing Company has a total cost function represented by the equation $y_c = 2x^3 - 3x^2 - 12x$, where y represents cost and x represents quantity.

- What equation represents the marginal cost function?
- What is the equation for the average cost function? At what point is average cost at its minimum?
- Is this a set of equations one might realistically expect to find in practice? Why?

6. The Minute Man Colonial Furniture Company's total revenue function is expressed by the equation $R = 24x - 3x^2$, where R represents revenue and x represents quantity.

- What is the maximum revenue the company can expect, assuming this equation is valid?
- What equation represents the average revenue function?
- What equation represents the marginal revenue function?
- On one graph, plot the total, average, and marginal revenue functions.

7. Arto Company manufactures cabinets for TV sets; the total cost of producing a certain model is given by $y = 4x - x^2 + 2x^3$, where y represents total cost and x represents quantity, in thousands of units. The sales department has indicated that between 2 and 6 units (in thousands) should be produced. At what quantity is marginal cost a minimum? Explain your answer, and graph the marginal cost curve.

8. If the general formula for the total cost function is

$$TC = f(x) = ax^3 + bx^2 + cx + d$$

- What is the corresponding equation for the marginal cost function?
- What is the corresponding equation for the average cost function?

9. An important generalization is often made about the relationships between the total revenue (TR), average revenue (AR), and marginal revenue (MR) functions:

When $MR = 0$, TR is at a maximum point

When $MR > AR$, AR is increasing

When $MR < AR$, AR is decreasing

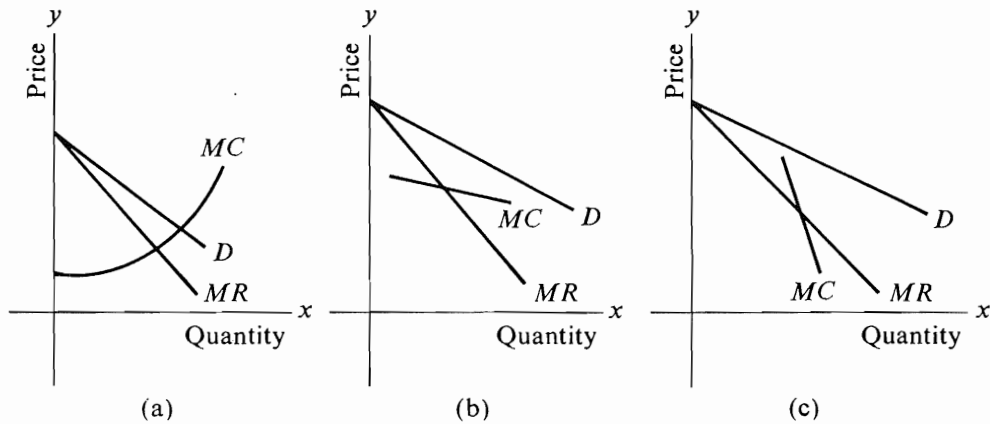
When $MR = AR$, AR is unchanging

- a. Illustrate this with an example from your reading or an example from the text.
 b. Work out the analogous relationships for the total cost (TC), average cost (AC), and marginal cost (MC) functions.
10. Henderson and Quandt in their book *Microeconomic Theory* state the following*:

"A monopolist is free to select any price-quantity combination which lies on his negatively sloped demand curve. Since an expansion of his output results in a reduction of his price, his MR is less than his price. His first-order condition for profit maximization requires the equality of MR and MC . His second-order condition requires that MC be increasing more rapidly than MR ." (Note that increasing more rapidly is logically equivalent to decreasing less rapidly.)

Which, if any, of the three diagrams in Figure 2.58 satisfies both conditions?

Figure 2.58



Answers to Odd-Numbered Problems

1. a. $\bar{y}_{\min} = 9$

b. $\bar{y}_{\min} = 2 - \frac{1}{e}$

c. $\bar{y}_{\min} = 2^{3/2}$

d. $\bar{y}_{\min} = 5 + 6\sqrt{2}$

e. $\bar{y}_{\min} = 17$

f. $\bar{y}_{\min} = 20$

g. $\bar{y}_{\min} = 9$

h. $\bar{y}_{\min} = 7 + 12\sqrt{6}$

3. a. for $0 \leq x \leq 10$, marginal and average cost always decreasing
 b. marginal cost decreasing for $x < 1$ and increasing for $x > 1$; average cost always decreasing; $\bar{y}_c \rightarrow 9$ as $x \rightarrow \infty$

* J. M. Henderson and R. E. Quandt, *Microeconomic Theory, a Mathematical Approach*, McGraw-Hill, New York, 1958, p. 198.

5. a. $MC = 6x^2 - 6x - 12$
 b. $AC = 2x^2 - 3x - 12$, min at $x = \frac{3}{4}$
 c. no, TC not in first quadrant for $x < \frac{3 + \sqrt{105}}{4}$
7. in interval $2 \leq x \leq 6$, MC is minimum for $x = 2$
9. b. when $MC = 0$, TC is maximum
 when $MC > AC$, AC is increasing
 when $MC < AC$, AC is decreasing
 when $MC = AC$, AC is unchanging

ELASTICITY

It is frequently of interest to consider the relative change in the dependent variable which results from a small relative change in the independent variable. The ratio of the relative change in the dependent variable, y , to the relative change in the independent variable, x , is referred to as the *elasticity* of y with respect to x . The elasticity of y with respect to x measures the responsiveness of y to changes in x . Because relative changes are used in the definition of elasticity, the elasticity of a function is dimensionless. That is, elasticity is independent of the units in which the variables are measured. Elasticities are frequently used in measuring the responsiveness of demand or supply to changes in price or income.

There are two types of measurement of elasticity, arc elasticity and point elasticity. *Arc elasticity* measures the elasticity of a function between two points, that is, over an arc. Unfortunately, the definition of arc elasticity is somewhat ambiguous, as discussed below. However, for a continuous function of a continuous variable, *point elasticity*, the elasticity of a function at a specific point, can be defined unambiguously.

If $y = f(x)$, the elasticity of y with respect to x is defined by

$$\frac{Ey}{Ex} = \frac{\Delta y}{y} \bigg/ \frac{\Delta x}{x} = \frac{x}{y} \cdot \frac{\Delta y}{\Delta x}$$

where, as $\Delta x \rightarrow 0$, $\frac{\Delta y}{\Delta x} \rightarrow \frac{dy}{dx}$. Elasticity is frequently denoted by η .

In general, the elasticity of a function varies over its range. For example, in Figure 2.59, note that elasticity of the function over the two arcs noted is not the same; although $\frac{\Delta y}{\Delta x} = \frac{\Delta y'}{\Delta x'}$, $\frac{x}{y} \neq \frac{x'}{y'}$ and thus $\frac{Ey}{Ex} \neq \frac{Ey'}{Ex'}$. Figure 2.59 also illustrates the ambiguity in defining arc elasticity since, for example, it is not clear whether x_1/y_1 or x_2/y_2 should be used for x/y .

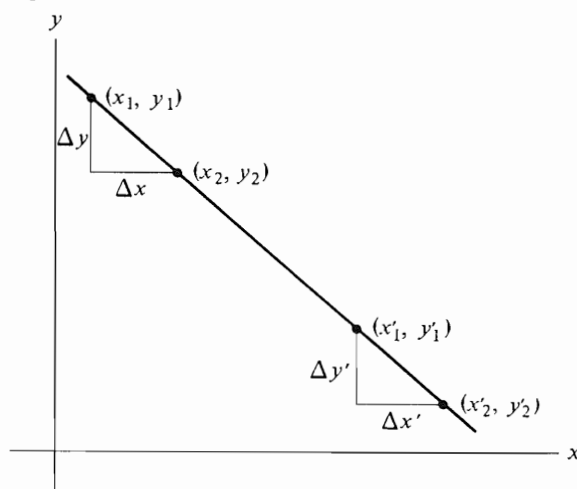
The following formulas are frequently used for measuring arc elasticity between the points (x_1, y_1) and (x_2, y_2) .

$$\frac{Ey}{Ex} = \frac{x_1}{y_1} \cdot \frac{y_2 - y_1}{x_2 - x_1} = \frac{x_1}{y_1} \cdot \frac{\Delta y}{\Delta x}$$

$$\frac{Ey}{Ex} = \frac{x_2}{y_2} \cdot \frac{y_2 - y_1}{x_2 - x_1} = \frac{x_2}{y_2} \cdot \frac{\Delta y}{\Delta x}$$

$$\frac{Ey}{Ex} = \frac{x_1 + x_2}{y_1 + y_2} \cdot \frac{y_2 - y_1}{x_2 - x_1} = \frac{x_1 + x_2}{y_1 + y_2} \cdot \frac{\Delta y}{\Delta x}$$

Figure 2.59



The first formula is an approximation of the point elasticity at (x_1, y_1) . The second formula is an approximation of the point elasticity at (x_2, y_2) . The third formula is the average of the first two and is probably the most commonly used.

For point elasticity, this ambiguity does not arise. The point elasticity of $y = f(x)$ at (x_1, y_1) is

$$\frac{Ey}{Ex} = \frac{x_1}{y_1} \cdot \left. \frac{dy}{dx} \right|_{x_1, y_1}$$

and, in general, the point elasticity of y with respect to x is given by

$$\eta = \frac{Ey}{Ex} = \frac{\frac{dy}{dx}}{\frac{y}{x}} = \frac{x}{y} \cdot \frac{dy}{dx}$$

The ratio of the derivative of a function to the function itself is called the *logarithmic derivative* of the function; that is, if $y = f(x)$, then

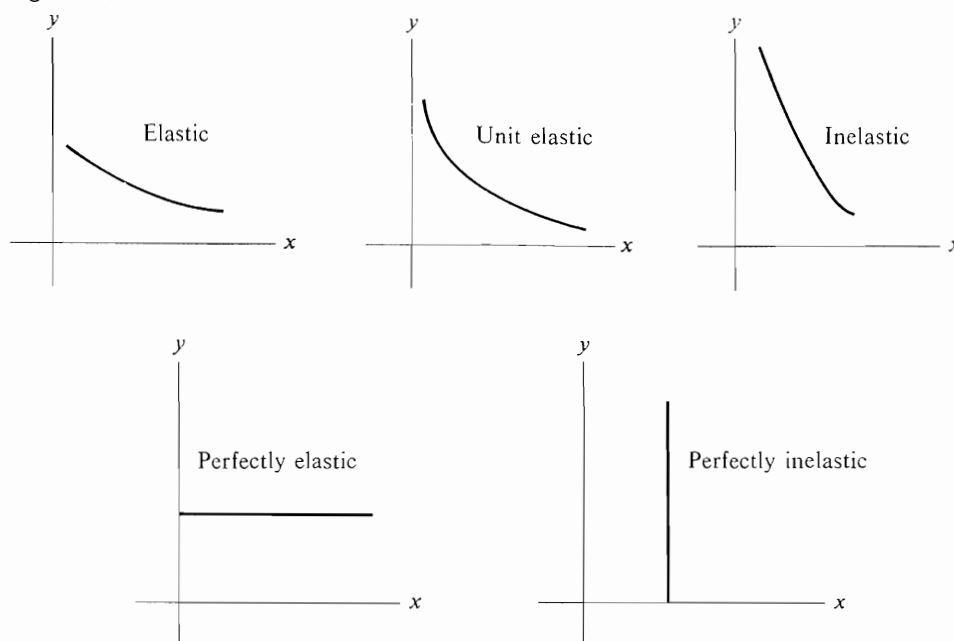
$$\frac{f'(x)}{f(x)} = \frac{\frac{dy}{dx}}{y} = \frac{1}{y} \frac{dy}{dx} = \frac{d}{dx} \ln y$$

is the logarithmic derivative of $f(x)$. The elasticity of the function $y = f(x)$ at the point x is the ratio of the logarithmic derivative of y to the logarithmic derivative of x :

$$\frac{Ey}{Ex} = \frac{\frac{d}{dx} \ln y}{\frac{d}{dx} \ln x} = \frac{\frac{1}{y} \frac{dy}{dx}}{\frac{1}{x} \frac{dx}{dx}} = \frac{x}{y} \cdot \frac{dy}{dx}$$

Elasticity is defined as a property of any differentiable function; however, in economic theory elasticity is most frequently considered for demand, supply, price, cost, and revenue functions. Elasticity is used, for example, as a measure of the responsiveness of demand or supply to changes in price or income, and the responsiveness of price, total cost, or total revenue to changes in quantity.

Figure 2.60



The classical and by far most frequent use of the concept of elasticity is in analyzing the responsiveness of demand for a commodity to changes in its price. Because the slope of a demand curve is negative, its first derivative is negative, and thus $\eta \leq 0$. Since elasticity is dimensionless, the responsiveness of different commodity demands to price changes can be compared. For this purpose, demand is frequently classified in general categories as perfectly elastic ($\eta = -\infty$), relatively elastic ($\eta < -1$), unit elastic ($\eta = -1$), relatively inelastic ($-1 < \eta < 0$), and perfectly inelastic ($\eta = 0$).

The price elasticity of demand for a commodity is of interest to a firm; for example, if demand is price elastic at a given price, a price decrease raises quantity demanded proportionately more than the price decreases, so that total revenue, which is the product of price and quantity, increases. When demand is unit elastic, total revenue is unchanged by a decrease in price; when demand is inelastic, total revenue is decreased by an increase in price. Demand curves having differing elasticities are shown in Figure 2.60. Note, however, that elasticity is not generally constant over the range of a function; note also that slope is only one aspect of elasticity and should not be equated with elasticity.

The concept of *cross elasticity* is useful as a measure of the relationship between the demands for two or more commodities. Cross elasticity measures the responsiveness of demand for one commodity to changes in the price of another commodity. The elasticity of demand for A with respect to price of B is defined as

$$\frac{Ex_A}{Ey_B} = \frac{\frac{dx_A}{x_A}}{\frac{dy_B}{y_B}} = \frac{y_B}{x_A} \cdot \frac{dx_A}{dy_B}$$

and *arc cross elasticity* as y_B changes from y_{B1} to y_{B2} and x_A changes from x_{A1} to x_{A2} is given by

$$\frac{Ex_A}{Ey_B} = \frac{y_{B1} + y_{B2}}{x_{A1} + x_{A2}} \cdot \frac{x_{A2} - x_{A1}}{y_{B2} - y_{B1}}$$

where the x_A 's are the quantities of A , the y_B 's are the prices of B , and the price of A is assumed to be constant. The elasticity of demand for B with respect to price of A is similarly defined.

When commodities are substitutes for each other, their cross elasticities are positive. For example, when the price of butter increases, the consumption of margarine also increases, assuming that the price of margarine and other relevant factors remain constant. When commodities are complementary to each other, their cross elasticities are negative. For example, when the price of lettuce increases, the consumption of salad dressing decreases, assuming that other relevant factors remain constant.

NOTE: Demand functions are customarily written with price as the dependent variable and quantity demanded as the independent variable and are graphed in a corresponding manner. However, in analyses involving demand elasticities, demand is logically considered to be the dependent variable and price the independent variable. In the examples below, the demand function is given in one form in some cases and in the other form in other cases and is represented graphically in both forms. When a demand function is given in the form $y = f(x)$ and it is not convenient to solve for x , $\frac{dx}{dy}$ can be obtained by implicit differentiation.

Because demand functions are conventionally written $y = f(x)$, where y is price and x is quantity demanded, and the elasticity of demand involves the derivative $\frac{dx}{dy}$, equations and graphs must be read carefully to avoid misinterpretation.

EXAMPLE

The demand function for a particular commodity is given by $y = \sqrt{(20 - x)/2}$ for $0 \leq x \leq 20$, where y is price per unit and x is number of units demanded. Consider the point $y = 2$, $x = 12$. If the price decreases by 6%, determine the corresponding increase in demand and an approximation to the elasticity of demand at the point $y = 2$, $x = 12$; compare this with the exact elasticity of demand at the point $y = 2$, $x = 12$. Then compute an approximation to the elasticity of demand and the exact elasticity of demand at the point representing the changed price and quantity; also compute the arc elasticity based on average price and quantity (see Figure 2.61).

Figure 2.61

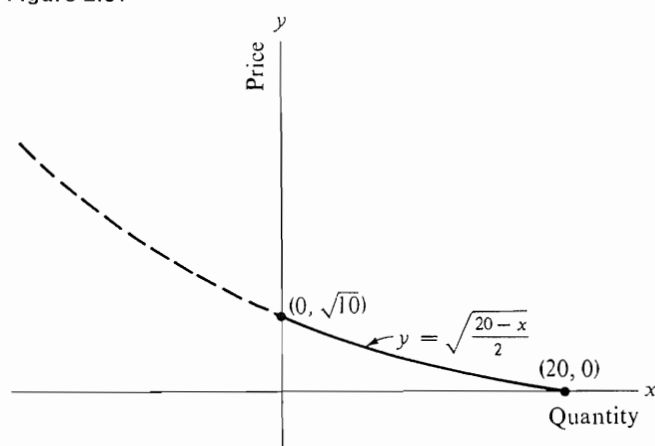
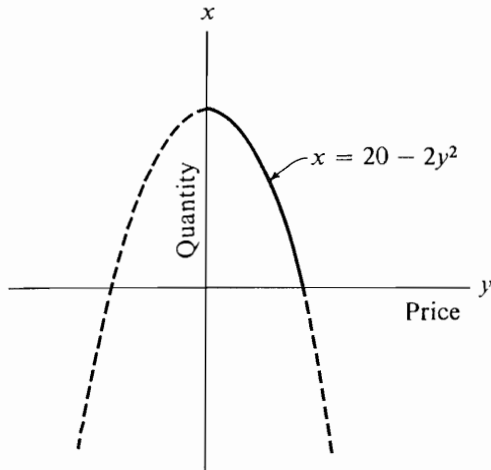


Figure 2.62



Solving the demand equation for x ,

$$y = \sqrt{\frac{20 - x}{2}}$$

$$y^2 = \frac{1}{2}(20 - x)$$

$$x = 20 - 2y^2$$

See Figure 2.62.

$$y_1 = 2 \quad y_2 = 2 - (0.06)(2) = 1.88 \quad \Delta y = -0.12$$

$$x_1 = 12 \quad x_2 = 20 - 2(1.88)^2 = 12.94 \quad \Delta x = 0.94$$

Arc elasticity of demand at $y = 2$, $x = 12$, is

$$\begin{aligned} \frac{Ex}{Ey} &= \frac{y_1}{x_1} \cdot \frac{\Delta x}{\Delta y} \\ &= \frac{2}{12} \left(-\frac{0.94}{0.12} \right) \\ &= -1.31 \end{aligned}$$

Point elasticity of demand at $y = 2$, $x = 12$, is

$$\begin{aligned} \frac{Ex}{Ey} &= \frac{y}{x} \cdot \frac{dx}{dy} \\ &= \frac{y}{20 - 2y^2} (-4y) \\ &= \frac{2y^2}{y^2 - 10} \end{aligned}$$

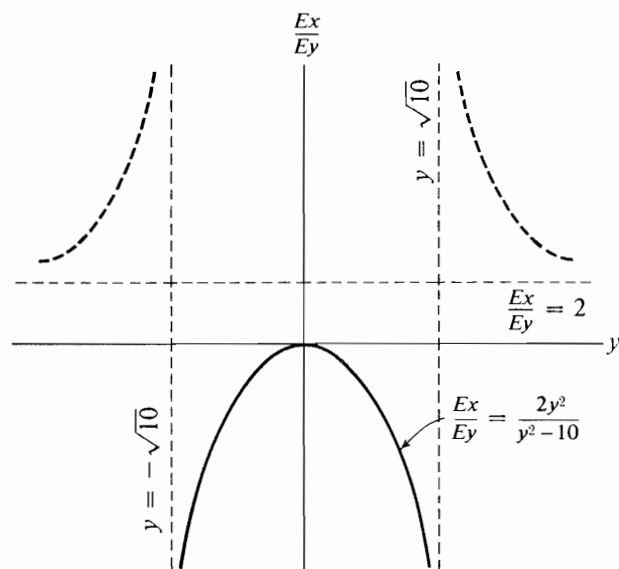
See Figure 2.63.

$$\left. \frac{Ex}{Ey} \right|_{y=2} = -\frac{8}{6} = -1.33$$

Arc elasticity of demand at $y_2 = 1.88$, $x_2 = 12.94$, is

$$\begin{aligned} \frac{Ex}{Ey} &= \frac{y_2}{x_2} \cdot \frac{\Delta x}{\Delta y} \\ &= \frac{1.88}{12.94} \left(-\frac{0.94}{0.12} \right) \\ &= -1.14 \end{aligned}$$

Figure 2.63



Point elasticity of demand at $y_2 = 1.88$, $x_2 = 12.9$, is

$$\frac{Ex}{Ey} = \frac{2y^2}{y^2 - 10}$$

$$\left. \frac{Ex}{Ey} \right|_{y=1.88} = -1.09$$

Arc elasticity based on average price and quantity is

$$\begin{aligned} \frac{Ex}{Ey} &= \frac{y_1 + y_2}{x_1 + x_2} \cdot \frac{x_2 - x_1}{y_2 - y_1} = \frac{3.88}{24.94} \left(-\frac{0.94}{0.12} \right) \\ &= -1.22 \end{aligned}$$

EXAMPLE

The demand for margarine is related to the price of butter as follows:

$$q = 5 + 2\sqrt{p}$$

where q is the demand for margarine (hundreds of pounds) and p is the price of butter (dollars per pound). Find the equation for cross elasticity of demand for margarine with respect to price of butter. What is the demand for margarine when butter is \$0.81 per pound? What is the demand for margarine when butter is \$1.00 per pound? What is the cross elasticity of demand for margarine as butter goes from \$0.81 to \$1.00 per pound? (See Figure 2.64.)

Solving the demand equation for p ,

$$q = 5 + 2\sqrt{p}$$

$$\sqrt{p} = \frac{1}{2}(q - 5)$$

$$p = \frac{1}{4}(q^2 - 10q + 25)$$

See Figure 2.65.

Figure 2.64

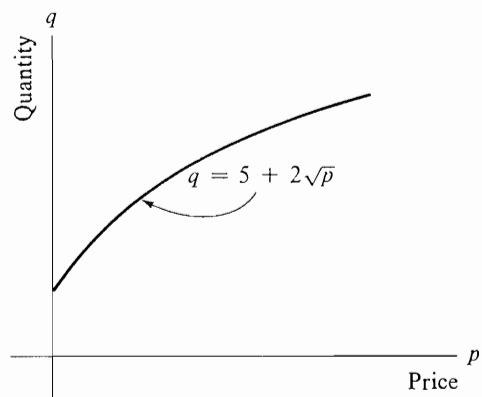
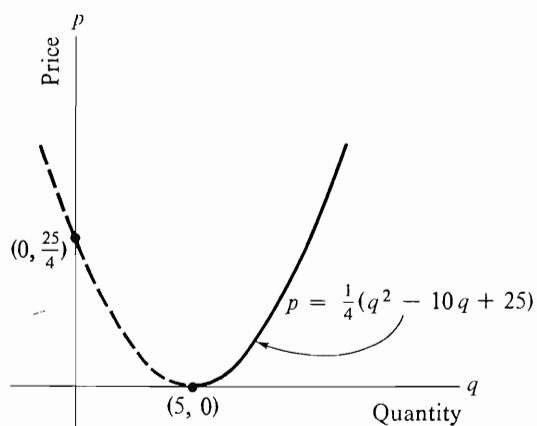


Figure 2.65

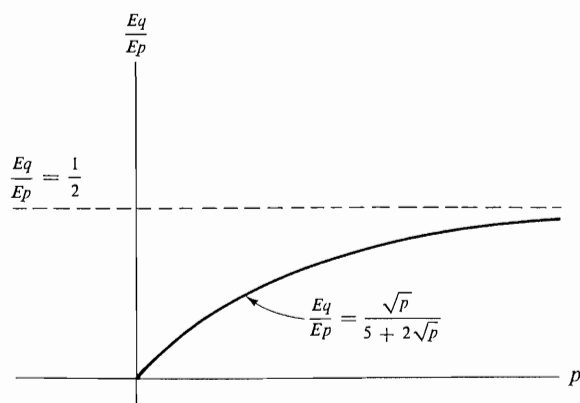


The elasticity of demand for margarine with respect to price of butter is

$$\begin{aligned} \frac{Eq}{Ep} &= \frac{dq}{dp} \cdot \frac{p}{q} = \frac{1}{\sqrt{p}} \cdot \frac{p}{5 + 2\sqrt{p}} \\ &= \frac{\sqrt{p}}{5 + 2\sqrt{p}} \end{aligned}$$

See Figure 2.66.

Figure 2.66



When $p = 0.81$, $q = 5 + 2\sqrt{.81}$ or 680 pounds.

When $p = 1.00$, $q = 5 + 2\sqrt{1.00}$ or 700 pounds.

Cross elasticity is

$$\frac{Eq}{Ep} = \frac{\sqrt{p}}{5 + 2\sqrt{p}}$$

For $p = 0.81$,

$$\left. \frac{Eq}{Ep} \right|_{p=0.81} = \frac{0.9}{5 + 2(0.9)} = \frac{0.9}{6.8} = 0.13$$

For $p = 1.00$,

$$\left. \frac{Eq}{Ep} \right|_{p=1.00} = \frac{1}{5 + 2} = \frac{1}{7} = 0.14$$

Constant Elasticity of Demand

As noted previously, the elasticity of a function is not, in general, constant over the range of the function. However, one type of function, the equilateral hyperbola, has constant elasticity over its range. In many theoretical analyses, the demand function is represented by an equilateral hyperbola, because it is convenient to have constant elasticity of demand over the range of the function.

If the demand function is given by the generalized equilateral hyperbola

$$x = \frac{a}{y^m}$$

then

$$\frac{dx}{dy} = -amy^{-m-1}$$

$$\begin{aligned} \frac{Ex}{Ey} &= \frac{y}{x} \cdot \frac{dx}{dy} \\ &= y \left(\frac{y^m}{a} \right) (-amy^{-m-1}) \\ &= -m \end{aligned}$$

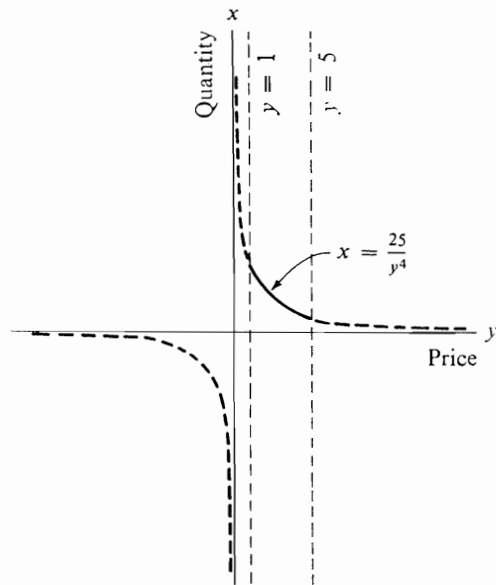
That is, elasticity of demand is the constant $-m$. In terms of approximate changes in price and quantity demanded, an increase of 1 percent in price results in a decrease of m percent in demand at any level of price and quantity. Only an equilateral hyperbola has constant elasticity over its range. For any other type of function, elasticity varies for different points on the curve. (See previous examples.)

EXAMPLE

The demand function is given by

$$x = \frac{25}{y^4} \quad \text{for } 1 \leq y \leq 5$$

Figure 2.67



Determine the elasticity of demand. Consider the point $y = 2$, $x = \frac{25}{16}$; if the price increases by 5 percent, determine arc elasticity at the point $y = 2$, $x = \frac{25}{16}$, and at the point representing the changed price and quantity (see Figure 2.67).

Solving the demand equation for y ,

$$x = \frac{25}{y^4}$$

$$y = \left(\frac{25}{x}\right)^{1/4}$$

See Figure 2.68.

Figure 2.68

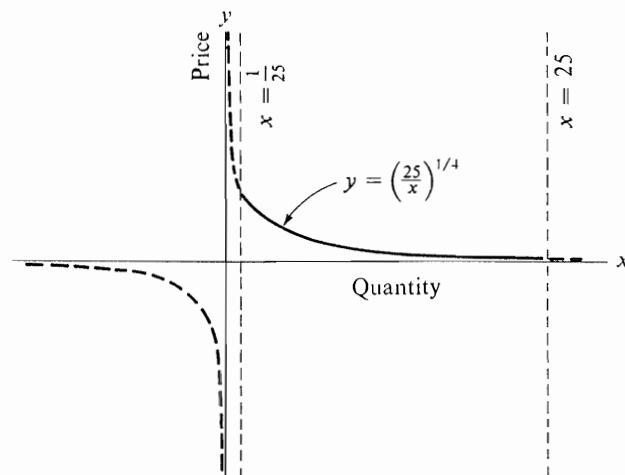
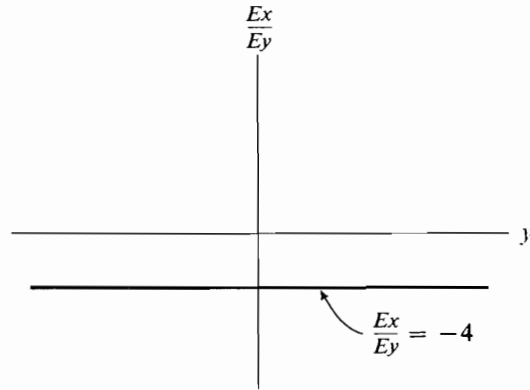


Figure 2.69



(Note that if $1 \leq y \leq 5$, then $\frac{1}{25} \leq x \leq 25$.)

$$\begin{aligned}\frac{dx}{dy} &= \frac{-100}{y^5} \\ \frac{Ex}{Ey} &= \frac{y}{x} \cdot \frac{dx}{dy} \\ &= \frac{y}{25} \left(\frac{-100}{y^5} \right) \\ &= -4\end{aligned}$$

as could be obtained by applying the derivation for the general case given above. See Figure 2.69.

$$\begin{aligned}y_1 &= 2 & y_2 &= 2 + (0.05)(2) = 2.1 & \Delta y &= 0.1 \\ x_1 &= 1.56 & x_2 &= \frac{25}{(2.1)^4} = 1.29 & \Delta x &= -0.27\end{aligned}$$

For the point $y_1 = 2$, $x_1 = 1.56$, arc elasticity of demand is

$$\begin{aligned}\frac{Ex}{Ey} &= \frac{y_1}{x_1} \cdot \frac{\Delta x}{\Delta y} \\ &= \frac{2}{1.56} \left(\frac{-0.27}{0.1} \right) \\ &= -3.46\end{aligned}$$

For the point $y_2 = 2.1$, $x_2 = 1.29$, arc elasticity of demand is

$$\begin{aligned}\frac{Ex}{Ey} &= \frac{y_2}{x_2} \cdot \frac{\Delta x}{\Delta y} \\ &= \frac{2.1}{1.29} \left(\frac{-0.27}{0.1} \right) \\ &= -4.40\end{aligned}$$

PROBLEMS

- For each of the following demand functions, i. determine arc elasticity at the point specified, ii. determine arc elasticity at the point corresponding to the specified change in price or

- demand, iii. determine point elasticity at the two points, iv. determine arc elasticity based on average quantities and prices, and v. show that $\frac{dy}{dx}$ and $\frac{dx}{dy}$ are reciprocals of each other.
- $x = 60 - 2y^2$; $x = 10$, $y = 5$; price decrease 8%
 - $x + 2y = 15$; $x = 7$, $y = 4$; price increase 5%
 - $x = 25 - 5y^2$; $x = 5$, $y = 2$; price increase 5%
 - $x = 10 - 5y^2$; $x = 5$, $y = 1$; price increase 10%
 - $y = (x - 10)^2$; $0 \leq x \leq 10$; $x = 8$, $y = 4$; demand decrease 5%
 - $x = 19 - 4y^2$; $x = 3$, $y = 2$; price decrease 5%
 - $x = 36 - 4y^2$; $x = 20$, $y = 2$; price increase 5%
 - $y = (x - 4)^2$; $x = 1$, $y = 9$; demand increase 30%
2. Find elasticity of demand, y , with respect to price, x , for each of the following functions.
- $y = \frac{3}{1 + 2x^2}$
 - $y = (x - 8)^2$, $0 \leq x \leq 8$
 - $y = ae^{-kx}$
 - $y = 13e^{-(5/4)x}$
 - $x = \frac{10}{y^{5/4}}$
 - $y = 100 - 15x^2$

Answers to Odd-Numbered Problems

- 9.6, -5.00, -10, -4.79, -6.73
 - 1.14, -1.27, -1.14, -1.27, -1.21
 - 8.2, -14.59, -8, -14.95, -10.57
 - 2.1, -2.92, -2, -3.06, -2.46
 - 0.11, -0.17, -0.13, -0.17, -0.14
 - 10.4, -6.5, -10.67, -6.33, -8.05
 - 1.64, -1.88, -1.6, -1.92, -1.75
 - 1.58, -0.98, -1.50, -1.04, -1.24

REVENUE, MARGINAL REVENUE, AND ELASTICITY OF DEMAND

For any given demand function $y = f(x)$, the total revenue R , is the product of x , the number of units demanded, and y , the price per unit quantity demanded,

$$R = xy = x \cdot f(x)$$

and the marginal revenue with respect to demand is

$$\frac{dR}{dx} = x \frac{dy}{dx} + y$$

Elasticity of demand with respect to price is

$$\frac{Ex}{Ey} = \frac{y}{x} \cdot \frac{dx}{dy}$$

and, since $\frac{dx}{dy} = \frac{1}{\frac{dy}{dx}}$,

$$\frac{Ex}{Ey} = \frac{y}{x} \left(\frac{1}{\frac{dy}{dx}} \right)$$

$$\frac{dy}{dx} = \frac{y}{x \left(\frac{Ex}{Ey} \right)}$$

Thus

$$\begin{aligned}\frac{dR}{dx} &= \frac{y}{\frac{Ex}{Ey}} + y \\ &= y \left(1 + \frac{1}{\frac{Ex}{Ey}} \right)\end{aligned}$$

Alternatively,

$$\begin{aligned}\frac{dR}{dx} &= x \frac{dy}{dx} + y \\ &= y \left(\frac{x}{y} \frac{dy}{dx} + 1 \right) \\ &= y \left(1 + \frac{Ey}{Ex} \right) \\ &= y \left(1 + \frac{1}{\frac{Ex}{Ey}} \right)\end{aligned}$$

That is, marginal revenue is the product of the price per unit quantity demanded and one plus the reciprocal of the elasticity of demand.

Since both x and y are either zero or positive, the total revenue, $R = xy$, is also either zero or positive; however, the marginal revenue may be either positive or negative.

NOTE: Logically, R could also be considered a function of price

$$R = xy = y \cdot g(y) = G(y)$$

and marginal revenue with respect to price defined as

$$\frac{dR}{dy} = y \frac{dx}{dy} + x$$

This would perhaps be simpler for cases when the demand function is in the form $x = g(y)$ and it is not easy to solve for y in terms of x . However, to keep the geometrical representation consistent (that is, to use x as the independent variable), it is preferable to use implicit differentiation and the chain rule to obtain $\frac{dR}{dx}$ and $\frac{d^2R}{dx^2}$ when the demand function is in the form $x = g(y)$.

EXAMPLE

The demand for a particular commodity is given by

$$y = (12 - x)^{1/2} \quad \text{for } 0 \leq x \leq 12$$

where x is quantity demanded and y is price per unit.

Determine the price and quantity for which revenue is maximum. Show for this demand function that the relationship between marginal revenue and elasticity of demand holds.

$$\text{Demand: } y = (12 - x)^{1/2} \quad \text{for } 0 \leq x \leq 12$$

Revenue: $R = x(12 - x)^{1/2}$

$$\frac{dR}{dx} = (12 - x)^{1/2} - \frac{1}{2}x(12 - x)^{-1/2}$$

$$= (12 - x)^{-1/2} [12 - x - \frac{1}{2}x]$$

$$= \frac{3(8 - x)}{2(12 - x)^{1/2}}$$

$$= 0 \quad \text{if } x = 8, R = 16$$

$$\frac{d^2R}{dx^2} = \frac{3}{2} \left[\frac{-(12 - x)^{1/2} - (8 - x)(-\frac{1}{2})(12 - x)^{-1/2}}{12 - x} \right]$$

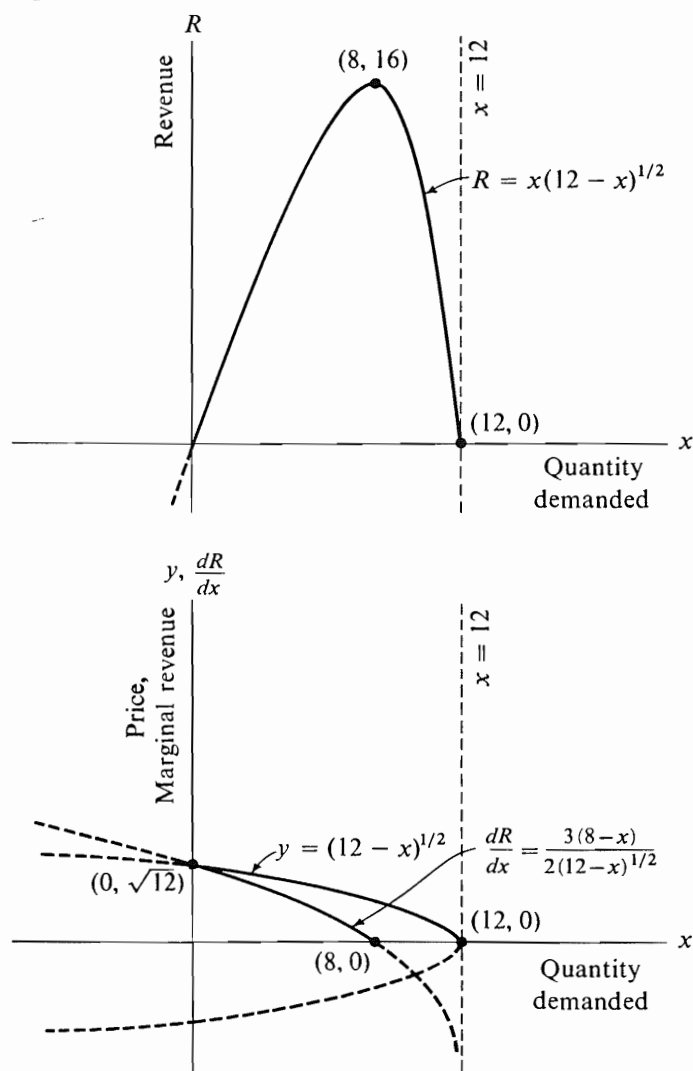
$$= \frac{3}{2} \left[\frac{(12 - x)^{-1/2} [-12 + x + \frac{1}{2}(8 - x)]}{12 - x} \right]$$

$$= \frac{3}{4} \left[\frac{x - 16}{(12 - x)^{3/2}} \right]$$

$$< 0 \quad \text{if } x = 8, \text{ so maximum at } x = 8, y = 16$$

$$< 0 \quad \text{for } 0 \leq x < 12, \text{ so no point of inflection}$$

Figure 2.70



Thus $R = x(12 - x)^{1/2}$ is concave downward for $0 \leq x \leq 12$. See Figure 2.70.

$$\begin{aligned}
 \frac{E_x}{E_y} &= \frac{y}{x} \cdot \frac{dx}{dy} \\
 &= \frac{y}{x} \left(\frac{1}{\frac{dy}{dx}} \right) \\
 &= \frac{(12 - x)^{1/2}}{x} \left[\frac{1}{-\frac{1}{2}(12 - x)^{-1/2}} \right] \\
 &= \frac{-2(12 - x)}{x} \\
 \frac{dR}{dx} &= y \left(1 + \frac{1}{\frac{E_x}{E_y}} \right) \\
 &= (12 - x)^{1/2} \left[1 - \frac{x}{2(12 - x)} \right] \\
 &= \frac{(12 - x)^{1/2}(24 - 3x)}{2(12 - x)} \\
 &= \frac{3(8 - x)}{2(12 - x)^{1/2}} \quad (\text{as above})
 \end{aligned}$$

EXAMPLE

The demand for a particular commodity is given by

$$y = 15e^{-x/3} \quad \text{for } 0 \leq x \leq 8$$

where x is quantity demanded and y is price per unit.

Determine the price and quantity for which revenue is maximum. Show for this demand function that the relationship between marginal revenue and elasticity of demand stated on page 267 holds.

$$\text{Demand: } y = 15e^{-x/3} \quad \text{for } 0 \leq x \leq 8$$

$$\text{Revenue: } R = 15xe^{-x/3}$$

$$\frac{dR}{dx} = -5xe^{-x/3} + 15e^{-x/3}$$

$$= 5e^{-x/3}(3 - x)$$

$$= 0 \quad \text{if } x = 3, R = \frac{45}{e}$$

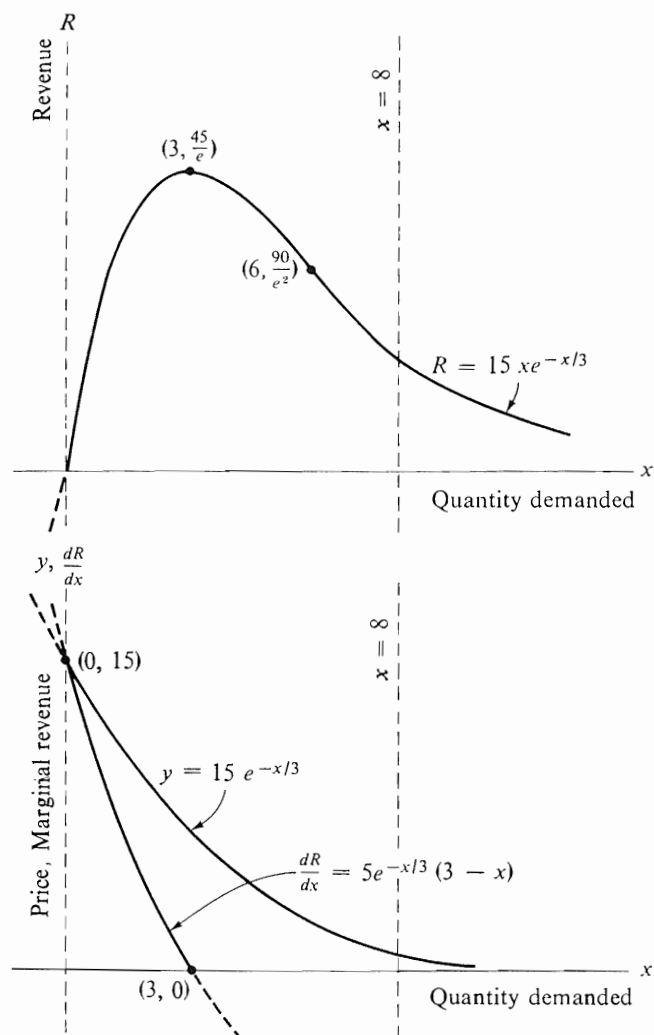
$$\frac{d^2R}{dx^2} = -\frac{5}{3}e^{-x/3}(3 - x) - 5e^{-x/3}$$

$$= 5e^{-x/3}\left(\frac{1}{3}x - 2\right)$$

$$< 0 \quad \text{if } x = 3, \text{ so maximum at } x = 3, R = \frac{45}{e} \approx 16.6$$

$$= 0 \quad \text{if } x = 6$$

Figure 2.71



$$\begin{aligned}\frac{d^3R}{dx^3} &= -\frac{5}{3}e^{-x/3}\left(\frac{1}{3}x - 2\right) + \frac{5}{3}e^{-x/3} \\ &= 5e^{-x/3}\left(1 - \frac{1}{3}x\right) > 0 \quad \text{if } x = 6\end{aligned}$$

so point of inflection at $x = 6$, $R = \frac{90}{e^2} \approx 12.2$. See Figure 2.71.

$$\begin{aligned}\frac{Ex}{Ey} &= \frac{y}{x} \cdot \frac{dx}{dy} \\ &= \frac{y}{x} \left(\frac{1}{\frac{dy}{dx}} \right) \\ &= \frac{15e^{-x/3}}{x} \left(\frac{1}{-5e^{-x/3}} \right) \\ &= -\frac{3}{x}\end{aligned}$$

$$\begin{aligned}
 \frac{dR}{dx} &= y \left(1 + \frac{1}{\frac{Ex}{Ey}} \right) \\
 &= 15e^{-x/3} \left(1 - \frac{x}{3} \right) \\
 &= 5e^{-x/3} (3 - x) \quad (\text{as above})
 \end{aligned}$$

PROBLEMS

1. When Mr. Smythe's income was \$300 per month he bought 20 qt of milk per month. When his income increased to \$350 per month he bought 24 qt of milk per month. Assuming no change in the price of milk or other relevant factors, what was Mr. Smythe's income elasticity of demand for milk?
2. When the price of A was \$5.00, 100 units of B were sold. When the price of A fell to \$4.00, 120 units of B were sold. What is the cross elasticity of demand for A in terms of the price of B ? Are A and B substitutes or complements?
3. What is the relation between the slope (positive or negative) of the average cost curve and the elasticity of total cost?
4. Demonstrate algebraically that the elasticity of total cost is equal to

$$E_{TC} = \frac{MC}{AC}$$

where MC is marginal cost and AC is average cost.

5. For each of the following demand functions demonstrate the relationship between marginal revenue and elasticity of demand given by

$$\frac{dR}{dx} = y \left(1 + \frac{1}{\frac{Ex}{Ey}} \right)$$

- a. $y = 550 - 3x - 6x^2$
 - b. $y = \frac{3250}{x^3}$
 - c. $y = 17 - 6x$
 - d. $y = 86 - 25x$
 - e. $y = 100 - 6x^2$
6. Show that if demand is linear and negatively sloped, the elasticity of total revenue is always less than 1, but to the left of its maximum value the elasticity of total revenue is positive and to the right of the maximum value the elasticity of total revenue is negative.

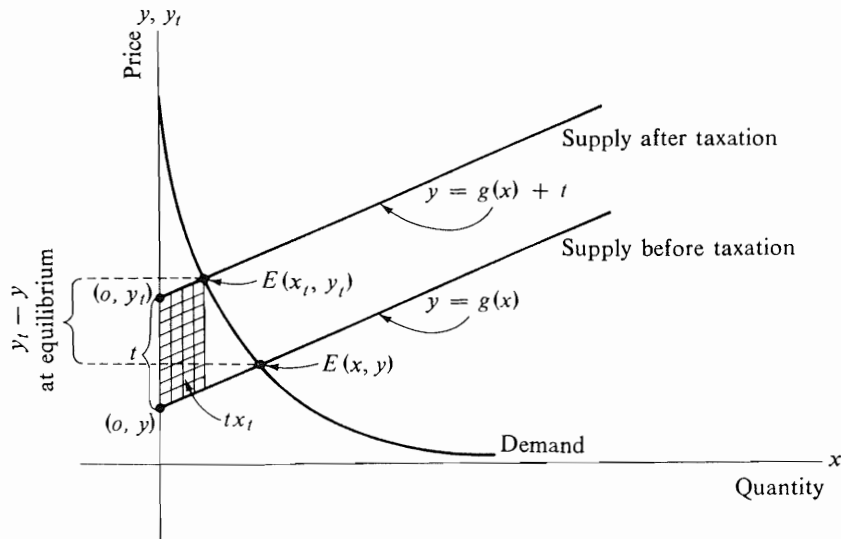
Answers to Odd-Numbered Problems

1. $\frac{13}{11}$
3. when slope of AC is negative, E_{TC} is less than 1
when slope of AC is zero, E_{TC} is equal to 1
when slope of AC is positive, E_{TC} is greater than 1

REVENUE FROM TAXATION

If the government imposes a tax on a given commodity, it is assumed that the price to the consumer will increase and correspondingly the quantity demanded will decrease. The discussion in this section concerns the effect of taxation on market

Figure 2.72



equilibrium under the following conditions: there is pure competition in which the consumers' demand depends only on price, the demand function does not change, the producers adjust the supply curve to the new price which includes the tax, and a tax of t monetary units is imposed on each unit of quantity produced.

The supply function can be represented by

$$y = g(x)$$

where x is the number of units of the commodity supplied and y is the price per unit. If a tax of t per unit quantity is imposed, the supply function after taxation is

$$y_t = g(x) + t$$

If the demand function is

$$y = f(x)$$

then the equilibrium point before taxation, $E(x, y)$, is the solution of the equations

$$\text{Demand: } y = f(x)$$

$$\text{Supply: } y = g(x)$$

and the equilibrium point after taxation, $E(x_t, y_t)$, is the solution of the equations

$$\text{Demand: } y = f(x)$$

$$\text{Supply: } y_t = g(x) + t$$

See Figure 2.72.

NOTE: If the supply function before taxation is in the form

$$x = G(y)$$

it may be possible to solve for y in a convenient form. If not, the supply function after taxation,

$$y_t - t = g(x)$$

shows that the quantity supplied is

$$x = G(y_t - t)$$

and equilibrium after taxation occurs at the solution of the equations

$$\text{Demand: } y = f(x)$$

$$\text{Supply: } x = G(y_t - t)$$

Geometrically, this is equivalent to moving the original supply curve upward t units. Except for the case of constant price, where price does not depend on quantity produced, the increase in equilibrium price is less than the amount of the tax. Note that a subsidy may be considered a negative tax. The supply curve is moved downward the amount of the subsidy, price to the consumer decreases, and demand increases.

The total tax revenue T received by the government from a tax of t per unit quantity is

$$T = tx_t$$

where x_t is the equilibrium quantity after taxation. This total may be represented by the area of a rectangle of dimensions t and x_t , which, if the supply function is linear, is equivalent to the parallelogram in Figure 2.72. As this diagram indicates, the tax may be large enough to reduce demand to the vanishing point. If the government is interested in taxation for revenue purposes, the question is what amount of tax will maximize revenue—if the tax t is considered as a variable, then there is no revenue if $t = 0$ and there is also no revenue if t is large enough so that demand is zero; for some intermediate value of t the tax revenue $T = tx_t$ is maximum. Since T may be considered a function of either t or x , maximum revenue possible from taxation can be determined by considering marginal tax revenue with respect to either t or x . It is usually more convenient to consider T as a function of x , since t occurs linearly in the relation between t and x .

EXAMPLE

The demand and supply functions for a particular commodity are

$$\text{Demand: } 2y + x = 14$$

$$\text{Supply: } y = \frac{3}{4} + \frac{x}{3}$$

Determine the maximum possible revenue from taxation which can be obtained from a tax of t per unit quantity and the corresponding tax rate.

The supply function after taxation is $y = \frac{3}{4} + (x/3) + t$. Thus, after taxation, at equilibrium,

$$y = 7 - \frac{x}{2} = \frac{3}{4} + \frac{x}{3} + t$$

$$t = \frac{25}{4} - \frac{5}{6}x$$

$$T = tx = \frac{25}{4}x - \frac{5}{6}x^2$$

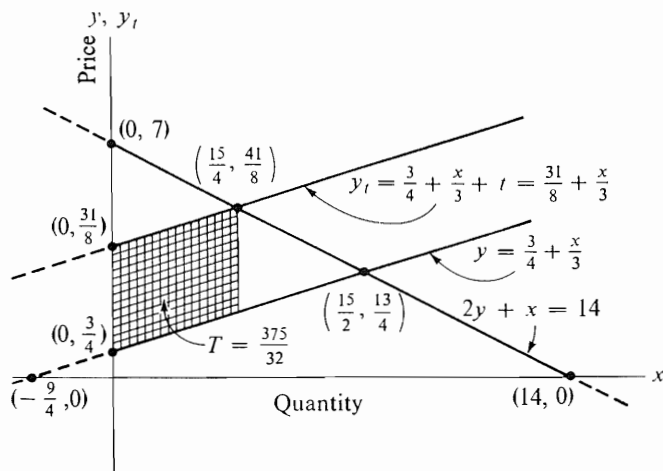
$$\frac{dT}{dx} = \frac{25}{4} - \frac{5}{3}x$$

$$= 0 \quad \text{if } x = \frac{15}{4}$$

$$\frac{d^2T}{dx^2} = -\frac{5}{3} < 0, \text{ so maximum at } x = \frac{15}{4}$$

If $x = \frac{15}{4}$, $t = \frac{25}{4} - \frac{5}{6}(\frac{15}{4}) = \frac{25}{8}$, and $T_{\max} = \frac{375}{32}$. So the maximum possible revenue from taxation is $T_{\max} = \frac{375}{32}$ obtained from a tax of $t = \frac{25}{8}$ per unit quantity. (See Figure 2.73.)

Figure 2.73



Alternatively,

$$x = \frac{15}{2} - \frac{6}{5}t$$

$$T = \frac{15}{2}t - \frac{6}{5}t^2$$

$$\frac{dT}{dt} = \frac{15}{2} - \frac{12}{5}t$$

$$= 0 \quad \text{if } t = \frac{25}{8}$$

$$\frac{d^2T}{dt^2} = -\frac{12}{5} < 0, \text{ so maximum at } t = \frac{25}{8}$$

$$T_{\max} = \left(\frac{15}{2}\right)\left(\frac{25}{8}\right) - \frac{6}{5}\left(\frac{25}{8}\right)^2 = \frac{375}{32}$$

EXAMPLE

The demand and supply functions for a particular commodity are

$$\text{Demand: } y = 30 - 2x^2$$

$$\text{Supply: } y = 3 + x^2$$

Determine the maximum possible revenue from taxation which can be obtained from a tax of t per unit quantity and the corresponding tax rate.

The supply function after taxation is $y = 3 + x^2 + t$. Thus, after taxation, at equilibrium,

$$y = 30 - 2x^2 = 3 + x^2 + t$$

$$t = 27 - 3x^2$$

$$T = tx = 27x - 3x^3$$

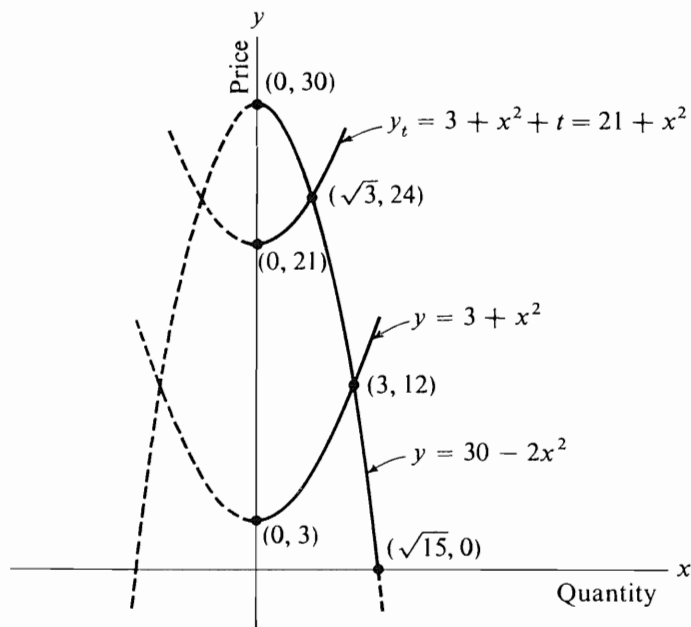
$$\frac{dT}{dx} = 27 - 9x^2$$

$$= 0 \quad \text{if } x = \sqrt{3}$$

$$\frac{d^2T}{dx^2} = -18x < 0 \quad \text{if } x = \sqrt{3}, \text{ so maximum at } x = \sqrt{3}$$

If $x = \sqrt{3}$, $t = 27 - 9 = 18$, and $T_{\max} = 18\sqrt{3}$. Thus the maximum possible revenue from taxation is $T_{\max} = 18\sqrt{3}$, obtained from a tax of $t = 18$ per unit quantity. (See Figure 2.74).

Figure 2.74



Alternatively,

$$x = \sqrt{\frac{27-t}{3}}$$

$$T = t \left(\frac{27-t}{3} \right)^{1/2} = \left(\frac{27t^2 - t^3}{3} \right)^{1/2}$$

$$\frac{dT}{dt} = \frac{1}{2} \left(\frac{27t^2 - t^3}{3} \right)^{-1/2} (54t - 3t^2)$$

$$= \frac{1}{2} \left(\frac{27-t}{3} \right)^{-1/2} (18-t)$$

$$= 0 \quad \text{if } t = 18$$

$$\frac{d^2T}{dt^2} = \left(-\frac{1}{2} \right) \left(\frac{1}{2} \right) \left(\frac{27-t}{3} \right)^{-3/2} (-1)(18-t) - \frac{1}{2} \left(\frac{27-t}{3} \right)^{-1/2}$$

$$= \left(\frac{27-t}{3} \right)^{-3/2} \left[\frac{1}{4}(18-t) - \frac{1}{2} \left(\frac{27-t}{3} \right) \right]$$

$$= -\frac{t}{12} \left(\frac{27-t}{3} \right)^{-3/2} < 0 \quad \text{if } t = 18, \text{ so maximum at } t = 18$$

$$T_{\max} = 18 \left(\frac{27-18}{3} \right)^{1/2} = 18\sqrt{3}$$

PROFIT UNDER MONOPOLY

In the usual forms of imperfect competition, it is assumed that the demand function $y = f(x)$ is known and that the price the consumer must pay depends only on the quantity demanded. In a monopolistic situation, the monopolist controls the price by regulating the supply of the commodity—when supply is limited, price is relatively high, and when supply increases, price decreases.

If \bar{y}_c is the average cost of producing a unit of commodity (as a function of the quantity produced), then the total cost y_c of producing x units is

$$y_c = x\bar{y}_c$$

(Note that in previous sections y has been used to denote price in some contexts and cost in others; the subscript c is used here to differentiate cost from price.)

Presumably the monopolist will control the supply x and thus the price y (determined by the demand function) so as to maximize his profit. The total revenue he receives is

$$R = xy$$

where $y = f(x)$; the total profit P is the difference between total revenue and total cost

$$P = R - y_c = xy - x\bar{y}_c$$

P has a relative maximum if and only if

$$\frac{dP}{dx} = 0 \quad \text{that is} \quad \frac{dR}{dx} = \frac{dy_c}{dx}$$

and

$$\frac{d^2P}{dx^2} < 0 \quad \text{that is} \quad \frac{d^2R}{dx^2} < \frac{d^2y_c}{dx^2}$$

In order to be meaningful this maximum must occur in the interval for which the cost and demand functions have economic significance.

NOTE: If the demand function is in the form

$$x = F(y)$$

and it is not possible or convenient to solve for y in terms of x , profit may be expressed as a function of y and the problem solved with y as the independent variable.

EXAMPLE

The demand function for a particular commodity is

$$y = 26 - 2x - 4x^2$$

and the average cost to the monopolist of producing and marketing the commodity is

$$\bar{y}_c = x + 8$$

Determine the maximum profit obtainable by the monopolist.

$$\text{Revenue: } R = xy = 26x - 2x^2 - 4x^3$$

$$\text{Total cost: } y_c = x\bar{y}_c = x^2 + 8x$$

$$\begin{aligned} \text{Profit: } P &= 26x - 2x^2 - 4x^3 - x^2 - 8x \\ &= 18x - 3x^2 - 4x^3 \end{aligned}$$

$$\frac{dP}{dx} = 18 - 6x - 12x^2$$

$$= 0 \quad \text{if } 2x^2 + x - 3 = 0$$

$$(2x + 3)(x - 1) = 0$$

$$x = 1$$

($x = -\frac{3}{2}$ is not meaningful for this problem.)

$$\frac{d^2P}{dx^2} = -6 - 24x$$

$$< 0 \quad \text{if } x = 1, \text{ so maximum at } x = 1$$

$$P_{\max} = 11$$

Alternatively,

$$\frac{dR}{dx} = 26 - 4x - 12x^2$$

$$\frac{dy_c}{dx} = 2x + 8$$

$$\frac{dR}{dx} = \frac{dy_c}{dx} \quad \text{if } 26 - 4x - 12x^2 = 2x + 8$$

$$2x^2 + x - 3 = 0$$

$$(2x + 3)(x - 1) = 0$$

$$x = 1$$

$$\left. \begin{array}{l} \frac{d^2R}{dx^2} = -4 - 24x \\ \frac{d^2y_c}{dx^2} = 2 \end{array} \right\} \frac{d^2R}{dx^2} < \frac{d^2y_c}{dx^2} \quad \text{if } x = 1, \text{ so maximum at } x = 1$$

$$P_{\max} = 11$$

EXAMPLE

The demand function for a particular commodity is

$$y = 28 - 5x$$

and the total cost to the monopolist of producing and marketing the commodity is

$$y_c = x^2 + 4x$$

Determine the maximum profit obtainable by the monopolist.

$$\text{Revenue: } R = xy = 28x - 5x^2$$

$$\text{Total cost: } y_c = x\bar{y}_c = x^2 + 4x$$

$$\begin{aligned} \text{Profit: } P &= 28x - 5x^2 - x^2 - 4x \\ &= 24x - 6x^2 \end{aligned}$$

$$\frac{dP}{dx} = 24 - 12x$$

$$\frac{dP}{dx} = 0 \quad \text{if } 24 - 12x = 0$$

$$x = 2$$

$$\frac{d^2P}{dx^2} = -12 < 0, \text{ so maximum at } x = 2$$

$$P_{\max} = 24$$

Alternatively,

$$\frac{dR}{dx} = 28 - 10x$$

$$\frac{dy_c}{dx} = 2x + 4$$

$$\frac{dR}{dx} = \frac{dy_c}{dx} \quad \text{if } 28 - 10x = 2x + 4$$

$$x = 2$$

$$\left. \begin{array}{l} \frac{d^2R}{dx^2} = -10 \\ \frac{d^2y_c}{dx^2} = 2 \end{array} \right\} \frac{d^2R}{dx^2} < \frac{d^2y_c}{dx^2}, \text{ so maximum at } x = 2$$

$$P_{\max} = 24$$

or alternatively,

$$x = \frac{1}{5}(28 - y)$$

$$R = \frac{1}{5}y(28 - y)$$

$$\begin{aligned} y_c &= \frac{1}{25}(28 - y)^2 + \frac{4}{5}(28 - y) \\ &= \frac{1}{25}(28 - y)(28 - y + 20) \\ &= \frac{1}{25}(28 - y)(48 - y) \end{aligned}$$

$$\begin{aligned} \frac{dR}{dy} &= \frac{28}{5} - \frac{2}{5}y \\ &= -\frac{1}{25}(48 - y) - \frac{1}{25}(28 - y) \end{aligned}$$

$$\frac{dy_c}{dy} = \frac{2}{25}y - \frac{76}{25}$$

$$\begin{aligned} \frac{dR}{dy} &= \frac{dy_c}{dy} \quad \text{if } \frac{28}{5} - \frac{2}{5}y = \frac{2}{25}y - \frac{76}{25} \\ y &= 18 \end{aligned}$$

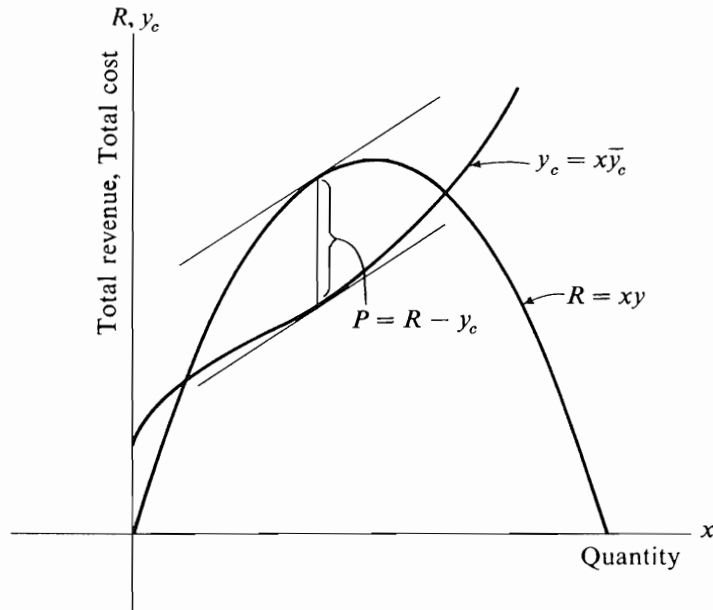
$$\left. \begin{array}{l} \frac{d^2R}{dy^2} = -\frac{2}{5} \\ \frac{d^2y_c}{dy^2} = \frac{2}{25} \end{array} \right\} \frac{d^2R}{dy^2} < \frac{d^2y_c}{dy^2}, \text{ so maximum at } y = 18$$

$$P_{\max} = R - y_c = 36 - 12 = 24$$

The following two geometric representations of profit under monopoly illustrate the economic relationships involved and also provide a method for determining approximate maximum profit in cases for which algebraic solution is difficult or impossible.

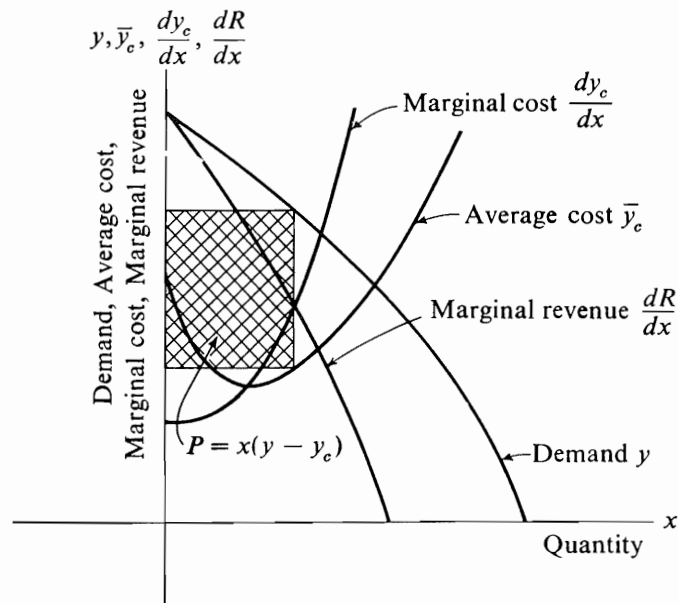
If the revenue and total cost curves are drawn on the same diagram, the vertical distance between them at any given value of x measures the profit corresponding to production of x units of the commodity. The maximum vertical distance corresponds to maximum profit and this occurs for the value of x where the slopes of the curves are equal, that is, where $\frac{dR}{dx} = \frac{dy_c}{dx}$ (see Figure 2.75).

Figure 2.75



If the marginal revenue and marginal cost curves are drawn on the same diagram, they intersect at a point which gives the value of x corresponding to maximum profit. If the demand curve $y = f(x)$ and the average cost curve $\bar{y}_c = g(x)$ are also drawn on the diagram, the vertical distance between them is $y - \bar{y}_c$. For any given value of x , the area of the rectangle of height $y - \bar{y}_c$ and width x represents profit; this area is greatest for the value of x where the marginal revenue and marginal cost curves intersect (see Figure 2.76).

Figure 2.76



EFFECT OF TAXATION ON MONOPOLY

The imposition of a tax t per unit quantity on a commodity produced by a monopolist increases the average cost by t and the total cost by tx . The equilibrium price and quantity for which the monopolist's profit is maximum after taxation are thus obtained by maximizing profit using the cost function after taxation,

$$y_{c_t} = y_c + tx$$

Profit after taxation is

$$P = R - y_{c_t} = R - y_c - tx$$

P has a relative maximum if and only if

$$\frac{dP}{dx} = 0 \quad \text{that is} \quad \frac{dR}{dx} = \frac{dy_{c_t}}{dx}$$

and

$$\frac{d^2P}{dx^2} < 0 \quad \text{that is} \quad \frac{d^2R}{dx^2} < \frac{d^2y_{c_t}}{dx^2}$$

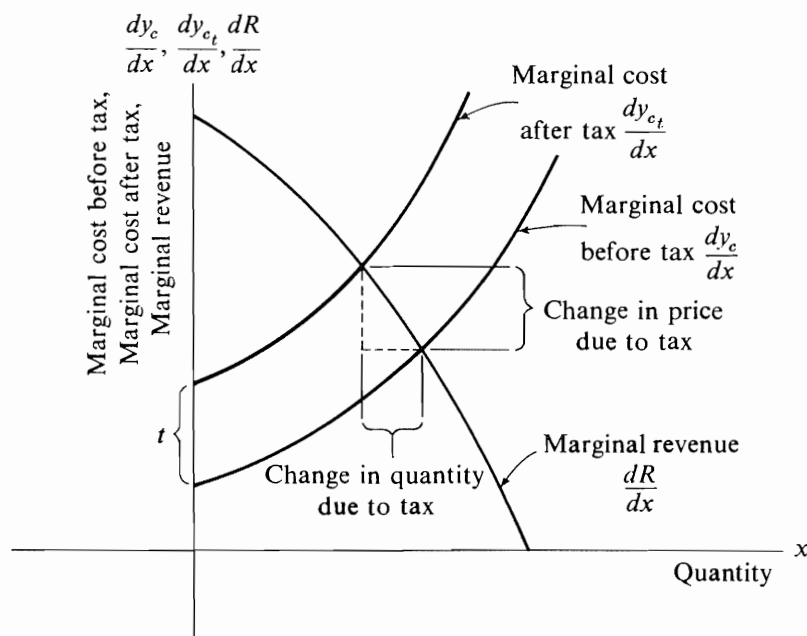
Since the marginal cost curve after taxation is the marginal cost curve before taxation translated upward a distance t , the amount produced for maximum profit will be decreased and the price increased after taxation. (See Figure 2.77).

The total revenue received by the government is

$$T = tx$$

where x is the amount produced after the tax is imposed. T is again a function of t and is zero if t is either zero or large enough to tax the product out of the market. Thus T has a maximum value obtainable in the usual way. As before, a subsidy may be considered a negative tax and the general analysis is unchanged.

Figure 2.77



EXAMPLE

The demand function for a particular commodity is

$$y = 20 - 4x$$

and the average cost to the monopolist is

$$\bar{y}_c = 2$$

- (a) If a tax of t per unit quantity is imposed on the monopolist, determine his maximum possible profit and the value of t for which tax revenue is maximized.
 (b) Determine the monopolist's maximum possible profit if a $33\frac{1}{3}\%$ sales tax is imposed.

$$\text{Demand: } y = 20 - 4x$$

$$\text{Revenue: } R = xy = x(20 - 4x)$$

$$\text{Average cost: } \bar{y}_c = 2$$

$$\text{Total cost: } y_c = 2x$$

$$\text{Total cost after tax: } y_{ct} = 2x + tx$$

$$\begin{aligned} \text{(a) Profit: } P &= R - y_{ct} = 20x - 4x^2 - 2x - tx \\ &= (18 - t)x - 4x^2 \end{aligned}$$

$$\frac{dP}{dx} = 18 - t - 8x$$

$$\frac{dP}{dx} = 0 \quad \text{if } x = \frac{18 - t}{8}, \quad y = \frac{22 + t}{2}$$

$$\frac{d^2P}{dx^2} = -8, \text{ so maximum if } x = \frac{18 - t}{8}$$

$$\begin{aligned} P_{\max} &= \frac{(18 - t)^2}{8} - 4\left(\frac{18 - t}{8}\right)^2 \\ &= \frac{(18 - t)^2}{16} \end{aligned}$$

$$\text{Tax revenue: } T = tx = \frac{t(18 - t)}{8}$$

$$\frac{dT}{dt} = \frac{18 - 2t}{8} = \frac{9 - t}{4}$$

$$= 0 \quad \text{if } t = 9$$

$$\frac{d^2T}{dt^2} = -\frac{1}{4} < 0, \text{ so maximum if } t = 9$$

If $t = 9$,

$$P_{\max} = \frac{81}{16}$$

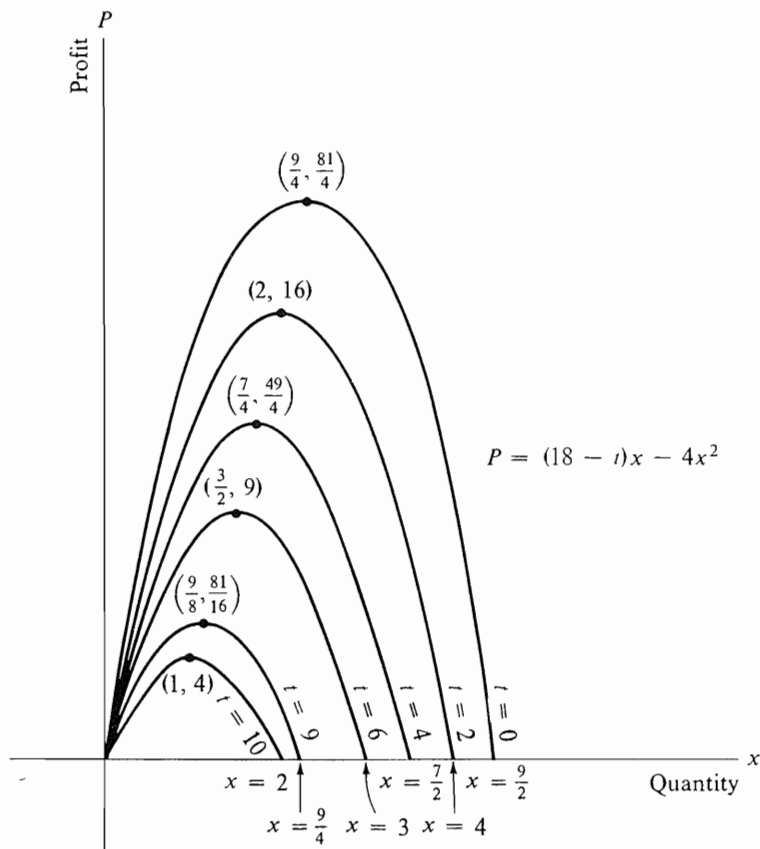
$$T_{\max} = \frac{81}{8}$$

See Figure 2.78.

(b) Demand function: $y = 20 - 4x$. So the price paid the monopolist after allowing for $33\frac{1}{3}\%$ sales tax is

$$y_t = \frac{3}{4}(20 - 4x)$$

Figure 2.78



Profit: $P = R - y_c = \frac{3}{4}x(20 - 4x) - 2x$

$$= 13x - 3x^2$$

$$\frac{dP}{dx} = 13 - 6x$$

$$= 0 \quad \text{if } x = \frac{13}{6}$$

$$\frac{d^2P}{dx^2} = -6, \text{ so maximum if } x = \frac{13}{6}$$

$$P_{\max} = \frac{169}{12}$$

EXAMPLE

The demand function for a particular commodity is

$$y = 14 - 3x$$

and the total cost to the monopolist is

$$y_c = x^2 + 5x$$

- (a) If a tax of t per unit quantity is imposed on the monopolist, determine his maximum possible profit, the change in price, and the tax revenue received by the government as a function of t .
- (b) Determine the maximum tax revenue obtainable by the government.

$$\text{Demand: } y = 14 - 3x$$

$$\text{Revenue: } R = xy = 14x - 3x^2$$

$$\text{Total cost: } y_c = x^2 + 5x$$

$$\text{Total cost after tax: } y_{c_t} = x^2 + 5x + tx$$

$$\begin{aligned} \text{(a) Profit: } P &= R - y_{c_t} = 14x - 3x^2 - x^2 - 5x - tx \\ &= (9 - t)x - 4x^2 \end{aligned}$$

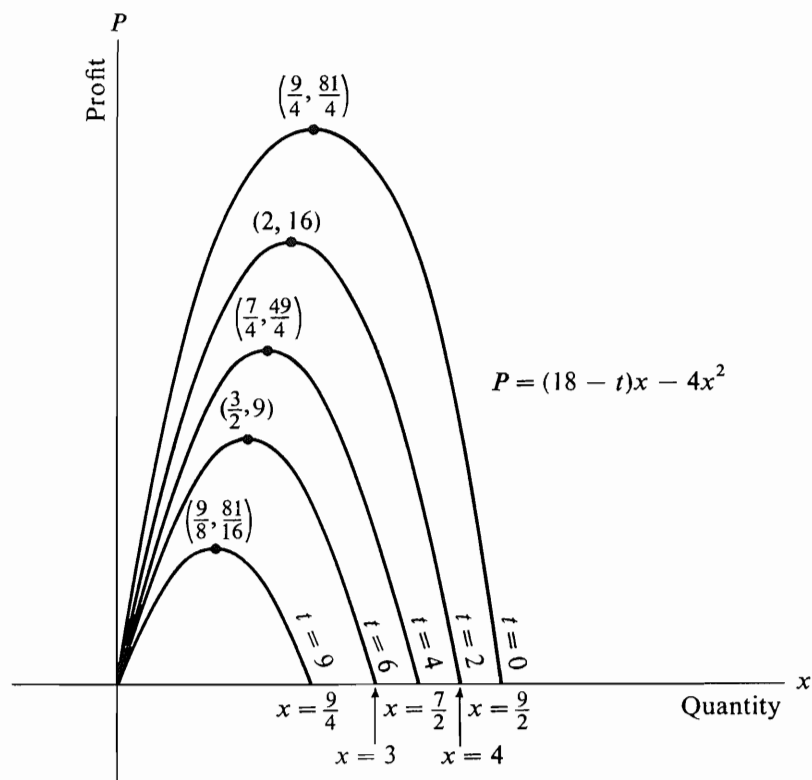
$$\frac{dP}{dx} = 9 - t - 8x$$

$$= 0 \quad \text{if } x = \frac{9-t}{8}, y = \frac{85}{8} + \frac{3}{8}t$$

$$\frac{d^2P}{dx^2} = -8, \text{ so maximum if } x = \frac{9-t}{8}$$

$$\begin{aligned} P_{\max} &= \frac{(9-t)^2}{8} - 4\left(\frac{9-t}{8}\right)^2 \\ &= \frac{(9-t)^2}{16} \end{aligned}$$

Figure 2.79



If $x = (9 - t)/8$, $y = \frac{85}{8} + \frac{3}{8}t$, which is a change in price of $\frac{3}{8}t$, since if $t = 0$ the monopolist's profit is maximized if $x = \frac{9}{8}$, $y = \frac{85}{8}$ (see Figure 2.79).

$$(b) \text{ Tax revenue: } T = tx = \frac{t(9 - t)}{8}$$

$$\frac{dT}{dt} = \frac{9}{8} - \frac{t}{4}$$

$$= 0 \quad \text{if } t = \frac{9}{2}$$

$$\frac{d^2T}{dt^2} = -\frac{1}{4}, \text{ so maximum if } t = \frac{9}{2}$$

$$T_{\max} = \frac{\frac{9}{2}(9 - \frac{9}{2})}{8} = \frac{81}{32}$$

PROBLEMS

- Plot the demand curve $y = 120 - x$ and the supply curve $y = x - 10$, where x is quantity demanded and y is price per unit, and indicate the equilibrium price and quantity. a. Assume that a tax of \$15 per unit is imposed on the seller; draw the new supply curve and indicate the new equilibrium. b. Assume that instead of a tax, a subsidy of \$10 is paid to the producers; draw the appropriate supply curve and indicate the corresponding equilibrium.
- For each of the following pairs of demand and supply functions find the maximum revenue that can be obtained by imposing a tax of t per unit and illustrate graphically (by sketching the demand curve, the supply curves before and after taxation, and the revenue curve).

a. Demand: $y = 14 - 3x$	d. Demand: $y = 28 - x^2$
Supply: $y = 4 + 2x$	Supply: $y = 4 + x^2$
b. Demand: $y = 25 - 2x^2$	e. Demand: $y = 45 - x^2$
Supply: $y = 5 + x$	Supply: $y = 6 + 2x$
c. Demand: $y = 50 - 4x^2$	f. Demand: $y = 30 - 2x^2$
Supply: $y = 5 + x^2$	Supply: $y = 3 + x^2$
- For each of the following pairs of demand and (average or total) cost functions find the maximum profit obtainable by a monopolist.

a. $y = 24 - 7x$	d. $y = 12 - 5x$
$\bar{y}_c = 6 - x$	$\bar{y}_c = 4x + 6$
b. $y = 26 - 3x^2$	e. $y = 26 - 2x - 4x^2$
$y_c = 3x^2 + 2x + 14$	$\bar{y}_c = x + 8$
c. $y = 12 - 4x$	
$y_c = 8x - x^2$	
- For each of the following pairs of demand and cost functions find i. change in price, maximum profit, and maximum revenue from taxation as functions of t if an additive tax of t per unit is imposed on a monopolist and determine the maximum revenue that can be obtained by such taxation, and ii. find the maximum revenue that can be obtained by the specified sales tax.

a. $y = 50 - 6x$	d. $y = 33 - 5x^2$
$y_c = x^2 + 9x$	$\bar{y}_c = 3x^2$
Sales tax 20%	Sales tax 10%
b. $y = 25 - 2x^2$	e. $y = 20 - 4x$
$y_c = 3x$	$\bar{y}_c = 2$
Sales tax 25%	Sales tax $33\frac{1}{3}\%$
c. $y = 12 - 4x$	f. $y = 72 - 7x^2$
$\bar{y}_c = 2x$	$\bar{y}_c = x^2$
Sales tax $33\frac{1}{3}\%$	Sales tax 40%

Answers to Odd-Numbered Problems

3. a. $P_{\max} = \frac{27}{2}$
 b. $P_{\max} = \frac{50}{9}$
 c. $P_{\max} = \frac{4}{3}$
 d. $P_{\max} = 1$
 e. $P_{\max} = 11$

INVENTORY MODELS

The objective of inventory control is to minimize total inventory cost. Inventory costs are of three types: (1) cost of placing an order or starting a production run (setup cost); (2) cost of holding inventory, including cost of capital or interest and storage cost (carrying cost); and (3) cost of going short, including loss of goodwill (shortage cost). Inventory control attempts to balance the economies of large orders or large production runs against the cost of holding inventory and, in some models, against the cost of going short.

In practice, various situations occur with respect to the nature of demand, the procedures involved in ordering or production, the ease of storage, the seriousness of shortages, and so forth. All the inventory models discussed in this book (including two in this section and one in Chapter 3) assume that demand is known and uniform. This assumption may be appropriate, at least as an approximation, in many situations involving demand for an input of production; however, it is entirely unrealistic in many situations involving retail demand. For these latter situations a model assuming probabilistic demand is more appropriate.

The models in this book also assume that the setup cost, the per unit carrying cost, and the per unit shortage cost do not depend on the number of items involved. Neither of the two models discussed in this section permits shortages; each of them can be derived as a special case of a model that does permit shortages by letting the shortage cost become very large. The first model assumes that items enter inventory in batches; the second model assumes that items enter inventory continuously during the period of ordering or production.

Model A: uniform demand, no shortages, batch arrival

D = demand per period

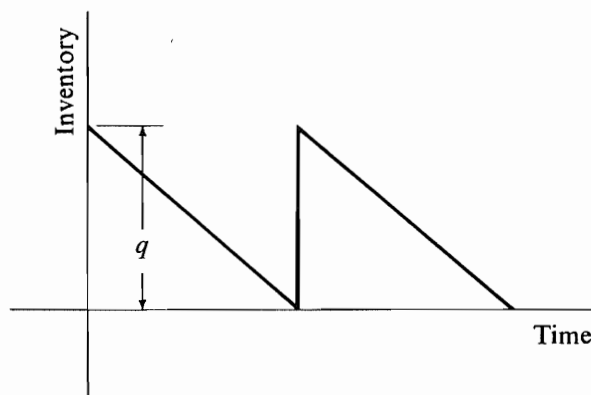
c_1 = setup cost

c_2 = carrying cost per item per period

q = number of items placed in inventory at one time

(See Figure 2.80.)

Figure 2.80



There are D/q batches per period for which the total setup cost is $(c_1 D)/q$. The average inventory is $q/2$ and the carrying cost per period is $(c_2 q)/2$. The total cost of inventory per period is thus

$$C = \frac{c_2 q}{2} + \frac{c_1 D}{q}$$

$$\frac{dC}{dq} = \frac{c_2}{2} - \frac{c_1 D}{q^2}$$

$$\text{If } \frac{dC}{dq} = 0, q = \sqrt{\frac{2c_1 D}{c_2}}.$$

$$\frac{d^2 C}{dq^2} = \frac{2c_1 D}{q^3} > 0$$

so C is minimum if $q = \sqrt{(2c_1 D)/c_2}$. Thus $\sqrt{(2c_1 D)/c_2}$ items should be placed in inventory D/q times per period.

Model B: uniform demand, no shortages, continuous arrival

D = demand per period

k = rate of arrival of items (number per period)

c_1 = setup cost

c_2 = carrying cost per item per period

q = number of items placed in inventory during buildup of inventory

t_1 = time during which items are placed in inventory

t_2 = time during which items are not placed in inventory

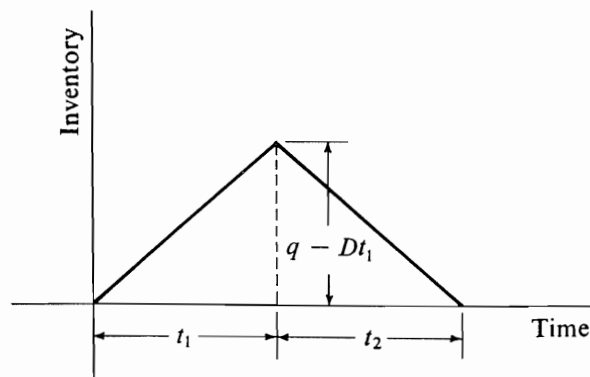
$t = t_1 + t_2$ = time required by one inventory cycle

(See Figure 2.81).

Note that $t_1 = q/k$ and $t = D/q$. The carrying cost per period is

$$\frac{c_2(q - Dt_1)}{2} = \frac{c_2 q \left(1 - \frac{D}{k}\right)}{2}$$

Figure 2.81



and the setup cost is $c_1(D/q)$. So the total inventory cost per period is

$$C = \frac{1}{2}c_2q\left(1 - \frac{D}{k}\right) + \frac{c_1D}{q}$$

$$\frac{dC}{dq} = \frac{1}{2}c_2\left(1 - \frac{D}{k}\right) - \frac{c_1D}{q^2}$$

$$= 0 \quad \text{if } q = \sqrt{\frac{2c_1D}{c_2\left(1 - \frac{D}{k}\right)}}$$

$$\frac{d^2C}{dq^2} = \frac{2c_1D}{q^3} > 0$$

So C is minimum if $q = \sqrt{(2c_1D)/c_2[1 - (D/k)]}$. Thus $\sqrt{(2c_1D)/c_2[1 - (D/k)]}$ items should be produced D/q times per period.

EXAMPLE

Alicor needs 2500 widgets per quarter. It costs \$3.00 a month to store a widget. The cost of ordering a supply of widgets (any number of widgets not exceeding 2500) is \$0.50. Alicor has two alternatives: (1) purchase the widgets periodically in lots and allow no shortages to occur and (2) purchase the widgets periodically from a supplier who sends them continuously at the rate of 1500 per month until an order is filled and allow no shortages to occur. Determine the quarterly cost of the optimal policy for each of these alternatives. In order to minimize total quarterly cost of inventory, what policy should Alicor adopt?

Considering demand and costs on a quarterly basis,

$$D = 2500$$

$$c_1 = 0.50$$

$$c_2 = 9.00$$

$$k = 4500$$

The optimal order quantity for alternative 1 is given by

$$q = \sqrt{\frac{2c_1D}{c_2}} = \sqrt{\frac{(2)(0.50)(2500)}{9}} = \frac{50}{3}$$

and the associated quarterly cost is

$$C = \frac{c_2q}{2} + \frac{c_1D}{q} = \frac{9(\frac{50}{3})}{2} + \frac{(\frac{1}{2})(2500)}{\frac{50}{3}} = 150$$

The optimal order quantity for alternative 2 is given by

$$q = \sqrt{\frac{2c_1D}{c_2\left(1 - \frac{D}{k}\right)}} = \sqrt{\frac{(2)(0.50)(2500)}{9\left(1 - \frac{2500}{4500}\right)}} = 25$$

and the associated quarterly cost is

$$C = \frac{1}{2}c_2q\left(1 - \frac{D}{k}\right) + \frac{c_1D}{q} = \frac{1}{2}(9)(25)\left(1 - \frac{25}{9}\right) + \frac{\frac{1}{2}(2500)}{25} = 100$$

Alternative 2 is preferable. As can be observed by comparing the cost formulas, this is always the case if the costs of ordering and carrying are the same for the two models.

For this problem, demand, supply and costs could be considered on a monthly basis. Since demand and supply are assumed to be uniform, any convenient common time base can be used for computation.

PROBLEMS

1. The Ace Manufacturing Company has a contract to supply 4000 refrigerators a year at a uniform monthly rate. The annual storage cost per refrigerator is \$50 and the setup cost for a production run is \$160. If production is instantaneous and shortages are not permitted, determine the number of refrigerators which should be produced in each run in order to minimize the total annual cost of inventory.
2. The Harvest Company has a contract to supply 500 tractors a month at a uniform daily rate. The monthly storage cost per tractor is \$10 and the setup cost for a production run is \$200. Production is at a constant rate of 1000 tractors per month. If shortages are not permitted, determine the number of tractors which should be produced in each run to minimize the total monthly inventory cost.
3. Profitco Electronics requires 25 transistors per month to make color TV sets. The monthly storage cost is \$2.00 per transistor and every time an order is placed it costs Profitco \$4.00.
 - a. If orders are filled immediately and no shortages are permitted, how many orders of what size should be placed every month? What is the monthly total cost of this inventory policy?
 - b. Suppose Profitco can order so that delivery is at the rate of 45 per month (in a uniform flow, not a block) and that no shortages are permitted. How much inventory should be accumulated before stopping the flow and how many times per month will the flow be stopped? What is the total monthly cost of this inventory policy?
4. Fabric Corporation owns a machine which can produce 25 bolts of a particular type of cloth per day and another machine which can make 10 bolts of this cloth into slipcovers per day. It costs \$1.50 to hold a bolt of cloth in inventory for a day and it costs \$2.00 every time the machine producing the cloth must be stopped.
 - a. What amount of inventory should be built up before the machine producing the cloth is stopped in order to minimize total inventory cost?
 - b. How many times will the machine be shut down in a period of 90 days?
 - c. What is the total amount of cloth that will be produced in a period of 90 days?
 - d. If the demand and cost conditions remain the same, except that the cost of stopping the machine is interpreted as the cost of ordering, and the cloth must be ordered in batches from another company, how many orders of what size will be placed in 90 days (assume no shortages)?
5. The yearly demand for a large cable bolt produced by Lanier Corporation is 9000. It costs Lanier \$1000 to put the production line for this cable into operation. The monthly cost of carrying this item in inventory is \$6000. Assuming that Lanier does not permit shortages to occur, how many cable bolts should be produced per production run and how frequently should production be stopped?
6. Bookworms Cooperative stocks a very popular gourmet cookbook for which there is a demand of 1500 copies per month. The publisher pays postage, but the cooperative pays \$2.50 for each order in secretarial and administrative costs. Carrying cost is \$0.50 per copy per month. In order to minimize total inventory cost, how many books should be ordered at one time? How frequently should orders be placed?

Answers to Odd-Numbered Problems

1. 160
3. a. 10 transistors $2\frac{1}{2}$ times per month
cost \$20
b. accumulate $\frac{20}{3}$ transistors and stop $\frac{5}{3}$ times per month
cost \approx \$13.33
5. 500 cable bolts and stop 18 times per year

2.13 Indeterminate Forms

In previous sections concerning limits, the evaluation of

$$\lim_{x \rightarrow a} f(x) \quad \text{when } f(a) = \frac{0}{0} \text{ or } \frac{\infty}{\infty}$$

is discussed for certain particular types of functions $f(x)$. In this section more general procedures for evaluating these and other indeterminate forms are discussed. These procedures are based on *L'Hospital's rule*:

$$\text{If } \begin{cases} f(a) = g(a) = 0 \\ \text{or} \\ f(a) = g(a) = \infty \\ \text{and} \\ \lim_{x \rightarrow a} \frac{f'(x)}{g'(x)} \text{ exists} \end{cases} \quad \text{then } \lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a} \frac{f'(x)}{g'(x)}$$

(this rule holds for finite or infinite a). If $\lim_{x \rightarrow a} \frac{f'(x)}{g'(x)}$ is itself an indeterminate form $\frac{0}{0}$ or $\frac{\infty}{\infty}$, then L'Hospital's rule is applied again and $\lim_{x \rightarrow a} \frac{f'(x)}{g'(x)} = \lim_{x \rightarrow a} \frac{f''(x)}{g''(x)}$, and so forth.

CAUTIONS

1. $\lim_{x \rightarrow a} \frac{f'(x)}{g'(x)} \neq \lim_{x \rightarrow a} \frac{d}{dx} \left(\frac{f(x)}{g(x)} \right)$, that is, the quotient of the derivatives is *not* in general equal to the derivative of the quotient.
2. $\lim_{x \rightarrow a} \frac{f'(x)}{g'(x)} = \lim_{x \rightarrow a} \frac{f''(x)}{g''(x)}$ if and *only if* $\lim_{x \rightarrow a} \frac{f'(x)}{g'(x)}$ is the indeterminate form $\frac{0}{0}$ or $\frac{\infty}{\infty}$, that is, L'Hospital's rule applies only to indeterminate forms; thus when the rule is applied successively, the first determinate answer is correct and, in general, further applications give incorrect answers.

EXAMPLE

$$\begin{aligned} &\text{Evaluate } \lim_{h \rightarrow 0} \frac{\sqrt{4+h}-2}{h}. \\ &\text{type } \frac{0}{0}, \text{ so } \lim_{h \rightarrow 0} \frac{\sqrt{4+h}-2}{h} = \lim_{h \rightarrow 0} \frac{\frac{1}{2}(4+h)^{-1/2}}{1} = \frac{1}{4} \end{aligned}$$

EXAMPLE

$$\begin{aligned} &\text{Evaluate } \lim_{x \rightarrow 0} \frac{e^x - (1+x)}{x^2}. \\ &\text{type } \frac{0}{0}, \text{ so } \lim_{x \rightarrow 0} \frac{e^x - (1+x)}{x^2} = \lim_{x \rightarrow 0} \frac{e^x - 1}{2x} \quad \left(\text{also type } \frac{0}{0} \right) \\ &= \lim_{x \rightarrow 0} \frac{e^x}{2} = \frac{1}{2} \end{aligned}$$

EXAMPLE

Evaluate $\lim_{x \rightarrow \infty} \frac{\ln x}{x}$.

$$\text{type } \frac{\infty}{\infty}, \text{ so } \lim_{x \rightarrow \infty} \frac{\ln x}{x} = \lim_{x \rightarrow \infty} \frac{1/x}{1} = 0$$

EXAMPLE

Evaluate $\lim_{x \rightarrow \infty} \frac{x^n}{e^x}$.

$$\begin{aligned} \text{type } \frac{\infty}{\infty}, \text{ so } \lim_{x \rightarrow \infty} \frac{x^n}{e^x} &= \lim_{x \rightarrow \infty} \frac{nx^{n-1}}{e^x} && \left(\text{also type } \frac{\infty}{\infty} \right) \\ &= \lim_{x \rightarrow \infty} \frac{n(n-1)x^{n-2}}{e^x} && \left(\text{also type } \frac{\infty}{\infty} \right) \\ &= \cdots = \lim_{x \rightarrow \infty} \frac{n!}{e^x} = 0 \end{aligned}$$

PROBLEMS

Evaluate the following limits.

1. $\lim_{y \rightarrow \infty} \frac{e^y}{y^2}$
2. $\lim_{x \rightarrow \infty} \frac{x}{e^x}$
3. $\lim_{x \rightarrow \infty} \frac{\ln x}{x^k} \quad (k > 0)$
4. $\lim_{x \rightarrow 0} \frac{e^x - e^{-x}}{\sin x}$
5. $\lim_{x \rightarrow 0} \frac{e^x - \ln(x+1) - 1}{x^2}$
6. $\lim_{x \rightarrow \pi/2} \frac{\sec x + 1}{\tan x}$
7. $\lim_{x \rightarrow \infty} \frac{x^2 + x - 1}{e^x + e^{-x}}$
8. $\lim_{x \rightarrow 0} \frac{e^x + e^{-x} - 2}{x^2}$
9. $\lim_{x \rightarrow \infty} \frac{x^k}{e^x} \quad (k > 0)$
10. $\lim_{x \rightarrow 1} \frac{\ln x}{x^2 - 1}$
11. $\lim_{x \rightarrow 1} \frac{\ln(1-x)}{\cot \pi x}$
12. $\lim_{x \rightarrow 0} \frac{\sin^2 x}{x}$
13. $\lim_{x \rightarrow 0} \frac{\sin x - x}{x^3}$
14. $\lim_{x \rightarrow \infty} \frac{x^k}{\ln x}$
15. $\lim_{x \rightarrow \infty} \frac{e^x}{x^k}$
16. $\lim_{x \rightarrow 0} \frac{e^x - e^{-x} - 2 \sin x}{3x^3}$
17. $\lim_{x \rightarrow 0} \frac{e^{2x} - 1}{x^2 - \sin x}$
18. $\lim_{x \rightarrow 0} \frac{2 - 3e^{-x} + e^{-2x}}{2x^2}$
19. $\lim_{x \rightarrow \infty} \frac{e^x}{x^3}$
20. $\lim_{x \rightarrow 2} \frac{x^4 - 4x^3 + 16}{x^3 - 8}$
21. $\lim_{x \rightarrow 0} \frac{4 - 3e^x - e^{-3x}}{4x^2}$
22. $\lim_{x \rightarrow 2} \frac{x^3 - 4x}{x^2 - 2x}$

- | | |
|--|---|
| 23. $\lim_{t \rightarrow 0} \frac{1 - \cos t - \frac{1}{2}t^2}{t^4}$ | 33. $\lim_{x \rightarrow 0} \frac{e^x - 1}{x^2 - x}$ |
| 24. $\lim_{x \rightarrow \infty} \frac{x^3}{e^x + x^2 - 2}$ | 34. $\lim_{x \rightarrow -3} \frac{x^3 + 27}{x + 3}$ |
| 25. $\lim_{h \rightarrow 0} \frac{(\sin h)/h - \cos h}{h^2}$ | 35. $\lim_{x \rightarrow \infty} \frac{x^2}{e^x - 1}$ |
| 26. $\lim_{x \rightarrow \infty} \frac{4x^2 + 3x - 6}{8x + 2}$ | 36. $\lim_{x \rightarrow 0} \frac{\cot x}{\cot 2x}$ |
| 27. $\lim_{y \rightarrow 0} \frac{4y^3 + 6y}{3y^3 + 2y}$ | 37. $\lim_{x \rightarrow 0} \frac{\tan x - x}{x - \sin x}$ |
| 28. $\lim_{x \rightarrow \infty} \frac{e^x}{x}$ | 38. $\lim_{x \rightarrow \pi/2} \frac{\ln \sin x}{(\pi - 2x)^2}$ |
| 29. $\lim_{z \rightarrow 0} \frac{\sin z^2 - \sin^2 z}{z^4}$ | 39. $\lim_{x \rightarrow \phi} \frac{\sin x - \sin \phi}{x - \phi}$ |
| 30. $\lim_{x \rightarrow \pi/4} \frac{1 - \sin 2x}{(\pi/4) - x}$ | 40. $\lim_{y \rightarrow 0} \frac{e^y + \sin y - 1}{\ln(1 + y)}$ |
| 31. $\lim_{x \rightarrow \pi} \frac{1 + \cos x}{(\pi - x)^2}$ | 41. $\lim_{x \rightarrow 0} \frac{\cot x}{\ln x}$ |
| 32. $\lim_{\theta \rightarrow 0} \frac{\tan \theta - \sin \theta}{\theta^2}$ | 42. $\lim_{x \rightarrow 0} \frac{\ln \sin 2x}{\ln \sin x}$ |

Answers to Odd-Numbered Problems

- | | |
|--------------------|---------------------|
| 1. ∞ | 23. $-\frac{1}{24}$ |
| 3. 0 | 25. $\frac{1}{3}$ |
| 5. 1 | 27. 3 |
| 7. 0 | 29. $\frac{1}{3}$ |
| 9. 0 | 31. $\frac{1}{2}$ |
| 11. 0 | 33. -1 |
| 13. $-\frac{1}{6}$ | 35. 0 |
| 15. ∞ | 37. 2 |
| 17. -2 | 39. $\cos \phi$ |
| 19. ∞ | 41. $-\infty$ |
| 21. $-\frac{3}{2}$ | |

L'Hospital's rule is also applicable to other types of indeterminate forms, if they are first put into the form $0/0$ or ∞/∞ , as follows.

Type $\infty \cdot 0$:

If $\lim_{x \rightarrow a} f(x)g(x) = \infty \cdot 0$, where $\lim_{x \rightarrow a} f(x) = \infty$, $\lim_{x \rightarrow a} g(x) = 0$, then $\lim_{x \rightarrow a} f(x)/[1/g(x)]$ is of type ∞/∞ or, alternatively, $g(x)/[1/f(x)]$ is of type $0/0$.

Type 0^0 :

If $\lim_{x \rightarrow a} f(x)^{g(x)} = 0^0$, where $\lim_{x \rightarrow a} f(x) = 0$, $\lim_{x \rightarrow a} g(x) = 0$, then

$$\lim_{x \rightarrow a} \left(\frac{g(x)}{\frac{1}{\log f(x)}} \right) \text{ is of type } \frac{0}{0}$$

or

$$\lim_{x \rightarrow a} \left(\frac{\log f(x)}{\frac{1}{g(x)}} \right) \text{ is of type } \frac{\infty}{\infty}$$

Type ∞^0 :If $\lim_{x \rightarrow a} f(x)^{g(x)} = \infty^0$, where $\lim_{x \rightarrow a} f(x) = \infty$, $\lim_{x \rightarrow a} g(x) = 0$, then

$$\lim_{x \rightarrow a} \left(\frac{g(x)}{\frac{1}{\log f(x)}} \right) \text{ is of type } \frac{0}{0}$$

or

$$\lim_{x \rightarrow a} \left(\frac{\log f(x)}{\frac{1}{g(x)}} \right) \text{ is of type } \frac{\infty}{\infty}$$

EXAMPLEEvaluate $\lim_{x \rightarrow \infty} x(e^{1/x} - 1)$.

$$\begin{aligned} \text{type } \infty \cdot 0, \text{ so } \lim_{x \rightarrow \infty} x(e^{1/x} - 1) &= \lim_{x \rightarrow \infty} \frac{e^{1/x} - 1}{\frac{1}{x}} \\ &= \lim_{x \rightarrow \infty} \frac{\frac{e^{1/x}}{x^2}}{-\frac{1}{x^2}} = 1 \end{aligned}$$

EXAMPLEEvaluate $\lim_{x \rightarrow 1} (1 - x) \tan \frac{1}{2}\pi x$.

$$\begin{aligned} \text{type } 0 \cdot \infty, \text{ so } \lim_{x \rightarrow 1} (1 - x) \tan \frac{1}{2}\pi x &= \lim_{x \rightarrow 1} \frac{1 - x}{\cot \frac{1}{2}\pi x} \\ &= \lim_{x \rightarrow 1} \frac{1}{-\frac{1}{2}\pi \csc^2 \frac{1}{2}\pi x} \\ &= \frac{1}{\pi/2} = \frac{2}{\pi} \end{aligned}$$

EXAMPLEEvaluate $\lim_{x \rightarrow 0} (1 - 2x)^{3/x}$.

type 1^∞ , so $\lim_{x \rightarrow 0} (1 - 2x)^{3/x} = \exp \lim_{x \rightarrow 0} \frac{\ln(1 - 2x)}{\frac{x}{3}}$

$$\lim_{x \rightarrow 0} \frac{\ln(1 - 2x)}{\frac{x}{3}} = \lim_{x \rightarrow 0} \frac{\frac{-2}{1 - 2x}}{\frac{1}{3}} = -6$$

$$\lim_{x \rightarrow 0} (1 - 2x)^{3/x} = e^{-6}$$

NOTE: Notation $\exp f(x)$ means $e^{f(x)}$.

EXAMPLE

Evaluate $\lim_{x \rightarrow 0} (\sin x + \cos x)^{1/x}$.

type 1^∞ , so $\lim_{x \rightarrow 0} (\sin x + \cos x)^{1/x} = \exp \lim_{x \rightarrow 0} \frac{\ln(\sin x + \cos x)}{x}$

$$\lim_{x \rightarrow 0} \frac{\ln(\sin x + \cos x)}{x} = \lim_{x \rightarrow 0} \frac{\frac{\cos x - \sin x}{\sin x + \cos x}}{1} = \frac{1}{1} = 1$$

$$\lim_{x \rightarrow 0} (\sin x + \cos x)^{1/x} = e^1 = e$$

EXAMPLE

Evaluate $\lim_{x \rightarrow 0^+} x^x$.

type 0^0 , so $\lim_{x \rightarrow 0^+} x^x = \exp \lim_{x \rightarrow 0^+} \frac{\ln x}{\frac{1}{x}}$

$$\lim_{x \rightarrow 0^+} \frac{\ln x}{\frac{1}{x}} = \lim_{x \rightarrow 0^+} \frac{\frac{1}{x}}{-\frac{1}{x^2}} = 0$$

$$\lim_{x \rightarrow 0^+} x^x = e^0 = 1$$

EXAMPLE

Evaluate $\lim_{x \rightarrow 1} (1 - x)^{\tan \pi x}$.

type 0^0 , so $\lim_{x \rightarrow 1} (1-x)^{\tan \pi x} = \exp \lim_{x \rightarrow 1} \frac{\ln(1-x)}{\cot \pi x}$

$$\begin{aligned} \lim_{x \rightarrow 1} \frac{\ln(1-x)}{\cot \pi x} &= \lim_{x \rightarrow 1} \frac{-1}{\frac{1-x}{\csc^2 \pi x}} \\ &= \lim_{x \rightarrow 1} \frac{\sin^2 \pi x}{x-1} \\ &= \lim_{x \rightarrow 1} (2 \sin \pi x \cos \pi x) \\ &= 0 \end{aligned}$$

$$\lim_{x \rightarrow 1} (1-x)^{\tan \pi x} = e^0 = 1$$

EXAMPLE

Evaluate $\lim_{x \rightarrow \infty} x^{1/x}$.

type ∞^0 , so $\lim_{x \rightarrow \infty} x^{1/x} = \exp \lim_{x \rightarrow \infty} \frac{\ln x}{x}$

$$\lim_{x \rightarrow \infty} \frac{\ln x}{x} = \lim_{x \rightarrow \infty} \frac{1/x}{1} = 0$$

$$\lim_{x \rightarrow \infty} x^{1/x} = e^0 = 1$$

EXAMPLE

Evaluate $\lim_{x \rightarrow \infty} (x + e^x)^{2/x}$

type ∞^0 , so $\lim_{x \rightarrow \infty} (x + e^x)^{2/x} = \exp \lim_{x \rightarrow \infty} \frac{\ln(x + e^x)}{\frac{x}{2}}$

$$\begin{aligned} \lim_{x \rightarrow \infty} \frac{\ln(x + e^x)}{\frac{x}{2}} &= \lim_{x \rightarrow \infty} \frac{\frac{1 + e^x}{x + e^x}}{\frac{1}{2}} \\ &= \lim_{x \rightarrow \infty} \frac{2(1 + e^x)}{x + e^x} \\ &= \lim_{x \rightarrow \infty} \frac{2e^x}{1 + e^x} \\ &= \lim_{x \rightarrow \infty} \frac{2e^x}{e^x} = 2 \end{aligned}$$

$$\lim_{x \rightarrow \infty} (x + e^x)^{2/x} = e^2$$

Type $\infty - \infty$:

If $\lim_{x \rightarrow a} [f(x) - g(x)] = \infty - \infty$, where $\lim_{x \rightarrow a} f(x) = \infty$, $\lim_{x \rightarrow a} g(x) = \infty$, then

$$\lim_{x \rightarrow a} \left(\frac{1}{\frac{1}{f(x)}} - \frac{1}{\frac{1}{g(x)}} \right) = \lim_{x \rightarrow a} \left(\frac{\frac{1}{g(x)} - \frac{1}{f(x)}}{\frac{1}{g(x)} \cdot \frac{1}{f(x)}} \right) \text{ is of type } \frac{0}{0}$$

EXAMPLE

Evaluate $\lim_{x \rightarrow 0} \left(\cot x - \frac{1}{x} \right)$.

$$\begin{aligned} \text{type } \infty - \infty, \text{ so } \lim_{x \rightarrow 0} \left(\cot x - \frac{1}{x} \right) &= \lim_{x \rightarrow 0} \frac{x - \tan x}{x \tan x} \\ &= \lim_{x \rightarrow 0} \frac{1 - \sec^2 x}{\tan x + x \sec^2 x} \\ &= \lim_{x \rightarrow 0} \frac{-2 \sec^2 x \tan x}{\sec^2 x + \sec^2 x + 2x \sec^2 x \tan x} \\ &= \lim_{x \rightarrow 0} \frac{-2 \tan x}{2(1 + x \tan x)} \\ &= \frac{0}{2} = 0 \end{aligned}$$

EXAMPLE

Evaluate $\lim_{x \rightarrow 0} \left(\frac{1}{x} - \frac{1}{e^x - 1} \right)$.

$$\begin{aligned} \text{type } \infty - \infty, \text{ so } \lim_{x \rightarrow 0} \left(\frac{1}{x} - \frac{1}{e^x - 1} \right) &= \lim_{x \rightarrow 0} \frac{e^x - 1 - x}{x(e^x - 1)} \\ &= \lim_{x \rightarrow 0} \frac{e^x - 1}{e^x + xe^x - 1} \\ &= \lim_{x \rightarrow 0} \frac{e^x}{e^x + e^x + xe^x} = \frac{1}{2} \end{aligned}$$

The types of indeterminate forms which can be evaluated using L'Hospital's rule are summarized in Table 2.1.

Table 2.1 Summary of Application of L'Hospital's Rule

Type of indeterminate form	Apply L'Hospital's rule to
1. $\lim \frac{f(x)}{g(x)} = \frac{0}{0}$	$\lim \frac{f(x)}{g(x)}$
2. $\lim \frac{f(x)}{g(x)} = \frac{\infty}{\infty}$	
3. $\lim f(x)g(x) = \infty \cdot 0$	$\lim \left(\frac{f(x)}{\frac{1}{g(x)}} \right)$ or to $\lim \left(\frac{g(x)}{\frac{1}{f(x)}} \right)$
4. $\lim f(x)^{g(x)} = 1^\infty$	$\lim \left(\frac{\ln f(x)}{\frac{1}{g(x)}} \right)$ or to $\lim \left(\frac{g(x)}{\frac{1}{\ln f(x)}} \right)$
5. $\lim f(x)^{g(x)} = 0^0$	If L'Hospital's rule applied to one of these gives a constant b for the limit, then the original limit is e^b
6. $\lim f(x)^{g(x)} = \infty^0$	
7. $\lim [f(x) - g(x)] = \infty - \infty$	$\lim \left(\frac{\frac{1}{g(x)} - \frac{1}{f(x)}}{\frac{1}{g(x)} \cdot \frac{1}{f(x)}} \right)$

PROBLEMS

Evaluate the following limits.

- $\lim_{x \rightarrow 1^-} x^{1/(1-x^2)}$
- $\lim_{x \rightarrow 0} x \cot x$
- $\lim_{x \rightarrow 0} x \ln \sin x$
- $\lim_{z \rightarrow \infty} z^{1/z}$
- $\lim_{x \rightarrow \pi/2} \tan x \ln \sin x$
- $\lim_{x \rightarrow 0} (1+x^2)^{1/x^2}$
- $\lim_{x \rightarrow 0} x^2 \csc x \cot x$
- $\lim_{x \rightarrow 0} x \ln x$
- $\lim_{x \rightarrow 0} (1 + \sin x)^{\cot x}$
- $\lim_{x \rightarrow \infty} \left(1 + \frac{2}{x}\right)^x$
- $\lim_{x \rightarrow \infty} x e^{-x}$
- $\lim_{x \rightarrow \pi/2} (\sec x - \tan x)$
- $\lim_{x \rightarrow \infty} (1+x)^{1/x}$
- $\lim_{x \rightarrow 0} x^n \ln x \quad (n > 0)$
- $\lim_{x \rightarrow 0} (e^x + x)^{1/x}$
- $\lim_{x \rightarrow \infty} (1+ax)^{b/x}$
- $\lim_{x \rightarrow 1^+} \left(\frac{1}{x-1} - \frac{1}{\sqrt{x-1}} \right)$
- $\lim_{x \rightarrow 1} \left(\frac{1}{\ln x} - \frac{x}{\ln x} \right)$
- $\lim_{x \rightarrow 1} \left(\frac{1}{\ln x} - \frac{x}{x-1} \right)$
- $\lim_{x \rightarrow \infty} (x-1)e^{-x^2}$
- $\lim_{x \rightarrow -\infty} x^2 e^x$
- $\lim_{x \rightarrow \pi/2} \sin x^{\tan x}$

23. $\lim_{x \rightarrow 0^-} (x + \sin x)^{\tan x}$
24. $\lim_{x \rightarrow \infty} (1 + e^{-x})^{e^x}$
25. $\lim_{x \rightarrow 0^+} x e^{1/x}$
26. $\lim_{x \rightarrow \infty} (x - \sqrt{x^2 + x})$
27. $\lim_{x \rightarrow 0} (e^x + 2x)^{1/x}$
28. $\lim_{x \rightarrow 1^-} x^{1/(1-x^2)}$
29. $\lim_{x \rightarrow 0} \frac{\sqrt{1-x} - \sqrt{1+x}}{x}$
30. $\lim_{x \rightarrow 0} \frac{1 - \sqrt{x+1}}{x}$
31. $\lim_{x \rightarrow 1} \csc \pi x \log x$
32. $\lim_{x \rightarrow \infty} (x-2)e^{-x^2}$
33. $\lim_{x \rightarrow \infty} x^2 e^{-3x}$
34. $\lim_{x \rightarrow 0^+} \sin x \log x$
35. $\lim_{\theta \rightarrow (\pi/2)^-} e^{-\tan \theta} \sec^2 \theta$
36. $\lim_{x \rightarrow 4} (x-4)^{x^2-16}$
37. $\lim_{x \rightarrow 1} x^{\cot \pi x}$
38. $\lim_{x \rightarrow 0^+} (x+1)^{\cot x}$
39. $\lim_{x \rightarrow \infty} (x^2 + a)^{1/x}$
40. $\lim_{x \rightarrow (\pi/2)^-} \tan x^{\cos x}$
41. $\lim_{x \rightarrow (\pi/2)^-} \sin x^{\tan x}$
42. $\lim_{x \rightarrow \infty} x^2 e^{-x}$
43. $\lim_{t \rightarrow 0} (1 + ct^2)^{c/t}$
44. $\lim_{x \rightarrow 0^+} \sin x^{\tan x}$
45. $\lim_{x \rightarrow 1} x^{(1/x^2)-1}$
46. $\lim_{y \rightarrow \infty} (e^y - 1)^{1/y}$
47. $\lim_{x \rightarrow 0^+} \csc x^{\sin x}$
48. $\lim_{x \rightarrow 0} x^2 \csc x$
49. $\lim_{\phi \rightarrow 0} \frac{\pi}{\phi} \tan \frac{\pi \phi}{2}$
50. $\lim_{x \rightarrow \infty} (1 + x^2)^{1/x}$
51. $\lim_{x \rightarrow \infty} x \sin \frac{a}{x}$
52. $\lim_{x \rightarrow (\pi/2)} (\pi - 2x) \tan x$
53. $\lim_{x \rightarrow 0} x^2 \ln x$
54. $\lim_{\theta \rightarrow (\pi/4)} (1 - \tan \theta) \sec 2\theta$
55. $\lim_{x \rightarrow 1} \left(\frac{1}{\ln x} - \frac{1}{x-1} \right)$
56. $\lim_{x \rightarrow 1^-} (\sqrt{2-x^2} - 1)^{x-1}$

Answers to Odd-Numbered Problems

- | | | |
|---------------|----------------------|-----------------------|
| 1. $e^{-1/2}$ | 21. 0 | 39. 1 |
| 3. 0 | 23. 0 | 41. 1 |
| 5. 0 | 25. 0 | 43. 1 |
| 7. 1 | 27. e^3 | 45. $e^{1/2}$ |
| 9. e | 29. -1 | 47. 1 |
| 11. 0 | 31. $-\frac{1}{\pi}$ | 49. $\frac{\pi^2}{2}$ |
| 13. 1 | 33. 0 | 51. a |
| 15. e^2 | 35. 0 | 53. 0 |
| 17. ∞ | 37. $e^{1/\pi}$ | 55. $\frac{1}{2}$ |

2.14 Sequences and Series

In this section, infinite sequences and series are defined and tests for determining convergence or divergence of an infinite series are considered. Power series are defined and the representation of a function by a Taylor's series is discussed.

A *sequence* is a succession of terms, u_1, u_2, \dots , formed according to some fixed rule; a *series* is the indicated sum $u_1 + u_2 + \dots$ of the terms of a sequence. A finite sequence or a finite series has a specified (finite) number of terms; an infinite sequence or an infinite series has an unspecified (infinite) number of terms. The general term, or n th term, of a sequence or series indicates the rule of formation of the terms.

For example, $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}$ is a finite sequence, and $1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5}$ is a finite series. The general term is $1/n$. However, if continued indefinitely, $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \frac{1}{7}, \frac{1}{8}, \dots$ is an infinite sequence and $1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} + \dots = \sum_{n=1}^{\infty} 1/n$ is an infinite series, still with the general term $1/n$.

Another method of specifying the terms of a sequence or series is by a *recursion formula* which gives the $(n+1)$ st term as a function of the preceding term or terms. For example, the infinite sequence 1, 2, 6, 24, 120, ... may be specified by the general term $u_n = n!$ or by the recursion formula $u_{n+1} = nu_n$, where u_n is the n th term and u_{n+1} is the $(n+1)$ st term. Many sequences are more conveniently specified by a recursion formula; other sequences, including the examples above, are as conveniently specified by a general term, but additional terms are more easily obtained by recursion. In order to specify a sequence completely by a recursion formula, the first term must be specified separately. Thus the sum of the sequence above is given by $\sum_{n=1}^{\infty} u_n = \sum_{n=1}^{\infty} 1/n!$ or by $\sum_{n=1}^{\infty} u_n = 1 + \sum_{n=2}^{\infty} nu_{n-1}$. There are some sequences for which it is necessary to specify the first term or terms separately, even if the general term is specified.

Although a finite sequence must have a finite sum, the limit of the sum of an infinite sequence may be finite or infinite.

Let S_n represent the sum of the first n terms of an infinite sequence $u_1, u_2, \dots, u_n, \dots$,

$$S_n = u_1 + u_2 + \dots + u_n = \sum_{i=1}^n u_i$$

and let S represent the limit of S_n as $n \rightarrow \infty$,

$$S = \lim_{n \rightarrow \infty} S_n = \sum_{i=1}^{\infty} u_i$$

If the limit S exists finitely, the infinite series is *convergent* and is said to converge to the value S ; if the limit does not exist finitely, the infinite series is *divergent*. Divergence may occur because S_n becomes infinite as $n \rightarrow \infty$ or because S_n oscillates without approaching a limit as $n \rightarrow \infty$.

Infinite series are essential in calculating the values of many functions and can also be used to define a number of useful functions. There are two problems generally associated with infinite series: determining whether the limit of an infinite series exists and determining the value of the limit, if it exists.

The problem of convergence is discussed before the use of infinite series in representing functions is considered. Establishing convergence or divergence of a series is relatively straightforward if an expression for S_n can be obtained.

EXAMPLE

For the geometric series of n terms,

$$S_n = a + ar + ar^2 + \dots + ar^{n-1}$$

it can be shown that

$$S_n = \frac{a(1 - r^n)}{1 - r} = \frac{a(r^n - 1)}{r - 1}$$

The first form is generally used if $|r| < 1$; the second form is generally used if $|r| > 1$.

If $|r| < 1$, $\lim_{n \rightarrow \infty} r^n = 0$ and $\lim_{n \rightarrow \infty} S_n = \frac{a}{1 - r}$ and the series is convergent.

If $|r| \geq 1$, $\lim_{n \rightarrow \infty} r^n = \infty$ and $\lim_{n \rightarrow \infty} S_n = \infty$ and the series is divergent.

NOTE: If $r = -1$, the geometric series is

$$a - a + a - a + a - a + \cdots$$

and if n is odd, $S_n = a$; if n is even, $S_n = 0$. Such a series has no limit and is thus divergent; it is referred to as an *oscillating series*.

Establishing convergence or divergence is clearly more difficult if an expression for S_n is not known. In such cases, the following tests are used.

1. *Necessary condition for convergence.* If an infinite series $\sum_{n=1}^{\infty} u_n$ is convergent, then $\lim_{n \rightarrow \infty} u_n = 0$. That is, if the n th term of a series does not approach 0 as n becomes infinite, the series is divergent. Note that $\lim_{n \rightarrow \infty} u_n = 0$ is a necessary but *not* a sufficient condition for convergence.
2. *Alternating-series test.* An alternating series is a series whose terms are alternately positive and negative. Such a series is convergent if $\lim_{n \rightarrow \infty} u_n = 0$ and each term is less in absolute value than the term which precedes it, that is, if $|u_{n+1}| < |u_n|$ for all $n = 1, 2, \dots$.

NOTE: It can be shown that the error involved in breaking a convergent alternating series at any term does not exceed in absolute value the first of the terms discarded. That is, if $u_1 - u_2 + u_3 - u_4 + \cdots$ is a convergent alternating series, then for any k ,

$$\left| \sum_{n=k+1}^{\infty} u_n \right| \leq |u_k|$$

3. *Absolute convergence.* A series of some positive and some negative terms is said to be *absolutely convergent* if the series formed from it by making all its terms positive is convergent. Other convergent series of some positive and some negative terms are said to be *conditionally convergent*. If a series of some positive and some negative terms is absolutely convergent, it is conditionally convergent, but the converse is not necessarily true.
4. *Cauchy's test-ratio test.* Let

$$u_1 + u_2 + u_3 + \cdots + u_{n+1} + \cdots$$

be an infinite series of positive terms. Using consecutive general terms u_n and u_{n+1} , form the test ratio

$$\frac{u_{n+1}}{u_n}$$

and let

$$\rho = \lim_{n \rightarrow \infty} \frac{u_{n+1}}{u_n}$$

Then

If $\rho < 1$, the series is convergent

If $\rho > 1$, the series is divergent

If $\rho = 1$, the test fails

5. *Comparison tests.* In many cases it is possible to determine whether a given series is convergent or divergent by comparing it term by term with a series known to be convergent or divergent. A series of positive terms is convergent if each of its terms is equal to or less than the corresponding term of a known convergent series. A series of positive terms is divergent if each of its terms is equal to or greater than the corresponding term of a known divergent series. The geometric series discussed above and the " p -series" are frequently useful in applying the comparison test. The " p -series" is given by

$$1 + \frac{1}{2^p} + \frac{1}{3^p} + \cdots + \frac{1}{n^p} + \cdots$$

This series is convergent if $p > 1$ and divergent if $p \leq 1$. If $p = 1$, the p -series is the harmonic series.

NOTE: Since convergence or divergence of a series cannot be affected by the omission of a finite number of terms, comparison tests may be applied to the terms $u_k, u_{k+1}, u_{k+2}, \dots$ rather than to the terms u_1, u_2, u_3, \dots .

SUMMARY OF PROCEDURE FOR TESTING CONVERGENCE OR DIVERGENCE OF AN INFINITE SERIES

In determining convergence or divergence of an infinite series, the tests should generally be applied in the order in which they are given above; thus the procedure is:

1. Determine whether $\lim_{n \rightarrow \infty} u_n$ is zero.

If $\lim_{n \rightarrow \infty} u_n \neq 0$, the series is divergent

If $\lim_{n \rightarrow \infty} u_n = 0$, further testing is necessary

Assuming $\lim_{n \rightarrow \infty} u_n = 0$,

2. Apply the alternating-series test (if appropriate): An alternating series whose terms decrease in numerical value and for which $\lim_{n \rightarrow \infty} u_n = 0$ is convergent.

3. Apply the ratio test: If $\lim_{n \rightarrow \infty} \frac{u_{n+1}}{u_n} = \rho$, then

$|\rho| < 1 \Rightarrow$ series convergent (absolutely)

$|\rho| > 1 \Rightarrow$ series divergent

$|\rho| = 1 \Rightarrow$ test fails

4. Apply the comparison test: Compare the series with a series known to be convergent or divergent, such as the geometric series or the p -series.

Geometric series: $a + ar + ar^2 + \cdots + ar^{n-1} + \cdots$

Convergent if $|r| < 1$

Divergent if $|r| \geq 1$

p -series: $1 + \frac{1}{2^p} + \frac{1}{3^p} + \cdots + \frac{1}{n^p} + \cdots$

Convergent if $p > 1$

Divergent if $p \leq 1$

EXAMPLE

Determine the convergence or divergence of the infinite series

$$\frac{1}{2} - \frac{3}{2^2} + \frac{5}{2^3} - \frac{7}{2^4} + \cdots$$

General term:

$$u_n = (-1)^{n+1} \frac{2n-1}{2^n}$$

$$\lim_{n \rightarrow \infty} u_n = 0$$

and $|u_{n+1}| < |u_n|$ for all n , so the alternating series is convergent.

$$\begin{aligned} \rho &= \lim_{n \rightarrow \infty} \frac{u_{n+1}}{u_n} = \lim_{n \rightarrow \infty} \left(-\frac{\frac{2(n+1)-1}{2^{n+1}}}{\frac{2n-1}{2^n}} \right) \\ &= \lim_{n \rightarrow \infty} \left(-\frac{2n+1}{2(2n-1)} \right) = \lim_{n \rightarrow \infty} \left(-\frac{1}{2} \right) = -\frac{1}{2} \\ |\rho| &= \left| -\frac{1}{2} \right| < 1 \end{aligned}$$

so the series is absolutely convergent.

EXAMPLE

Determine the convergence or divergence of the infinite series

$$\frac{2}{3} - \frac{3}{5} + \frac{4}{7} - \frac{5}{9} + \cdots$$

General term:

$$u_n = (-1)^{n+1} \frac{n+1}{2n+1}$$

$$\lim_{n \rightarrow \infty} u_n = \frac{1}{2}$$

so the series is divergent.

EXAMPLE

Determine the convergence or divergence of the infinite series

$$1 - \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{5}} - \frac{1}{\sqrt{7}} + \cdots$$

General term:

$$u_n = (-1)^{n+1} \frac{1}{\sqrt{2n-1}}$$

$$\lim_{n \rightarrow \infty} u_n = 0$$

and $|u_{n+1}| < |u_n|$ for all n , so the alternating series is convergent.

$$\begin{aligned}\rho &= \lim_{n \rightarrow \infty} \frac{u_{n+1}}{u_n} = \lim_{n \rightarrow \infty} \left(-\frac{\frac{1}{\sqrt{2(n+1)-1}}}{\frac{1}{\sqrt{2n-1}}} \right) \\ &= \lim_{n \rightarrow \infty} \left(-\frac{(2n-1)^{1/2}}{(2n+1)^{1/2}} \right) = -1 \\ |\rho| &= 1\end{aligned}$$

so the ratio test for absolute convergence fails.

$1/(2n-1)^{1/2} > (1/n)$ (p -series, $p = 1$), so the series is divergent as a positive series. Thus the alternating series is conditionally convergent.

EXAMPLE

Determine the convergence or divergence of the infinite series

$$\frac{1}{9} + \frac{2!}{9^2} + \frac{3!}{9^3} + \frac{4!}{9^4} + \cdots$$

General term:

$$u_n = \frac{n!}{9^n}$$

$$\lim_{n \rightarrow \infty} u_n = \infty$$

so the series is divergent.

EXAMPLE

Determine the convergence or divergence of the infinite series

$$\frac{3}{4} + 2\left(\frac{3}{4}\right)^2 + 3\left(\frac{3}{4}\right)^3 + 4\left(\frac{3}{4}\right)^4 + \cdots$$

General term:

$$u_n = n\left(\frac{3}{4}\right)^n$$

$$\lim_{n \rightarrow \infty} u_n = 0$$

$$\begin{aligned}\rho &= \lim_{n \rightarrow \infty} \frac{u_{n+1}}{u_n} = \lim_{n \rightarrow \infty} \left(\frac{(n+1)\left(\frac{3}{4}\right)^{n+1}}{n\left(\frac{3}{4}\right)^n} \right) \\ &= \lim_{n \rightarrow \infty} \frac{3}{4} \left(\frac{n+1}{n} \right) = \frac{3}{4}\end{aligned}$$

$$|\rho| = \left| \frac{3}{4} \right| < 1$$

so the series is convergent.

EXAMPLE

Determine the convergence or divergence of the infinite series

$$5 + \frac{5^2}{2!} + \frac{5^3}{3!} + \frac{5^4}{4!} + \cdots$$

General term:

$$u_n = \frac{5^n}{n!}$$

$$\lim_{n \rightarrow \infty} u_n = 0$$

$$\rho = \lim_{n \rightarrow \infty} \left(\frac{\frac{5^{n+1}}{(n+1)!}}{\frac{5^n}{n!}} \right) = \lim_{n \rightarrow \infty} \frac{5}{n+1} = 0$$

$$|\rho| = 0 < 1$$

so the series is convergent.

PROBLEMS

For each of the following series, write out the first few terms and determine whether the series is divergent or convergent (conditionally or absolutely for alternating series).

$$1. \sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{n^2 + 1}$$

$$2. \sum_{n=1}^{\infty} \frac{\sqrt{n}}{n^2 + 1}$$

$$3. \sum_{n=1}^{\infty} \frac{n}{n+2}$$

$$4. \sum_{n=1}^{\infty} \frac{1}{1 + \ln n}$$

$$5. \sum_{n=0}^{\infty} (-1)^{n+1} \frac{1}{(2n+1)!}$$

$$6. \sum_{n=1}^{\infty} \frac{(n+1)(n+2)}{n!}$$

$$7. \sum_{n=1}^{\infty} (-1)^{n+1} \frac{n^2}{2^n}$$

$$8. \sum_{n=0}^{\infty} \frac{(n+3)!}{3! n! 3^n}$$

$$9. \sum_{n=1}^{\infty} (-1)^{n+1} \frac{2^n}{n^3 + 1}$$

$$10. \sum_{n=1}^{\infty} \frac{10^n}{n!}$$

$$11. \sum_{n=1}^{\infty} \frac{n!}{9^n}$$

$$12. \sum_{n=1}^{\infty} \frac{n^2}{10(2n-1)}$$

$$13. \sum_{n=1}^{\infty} e^{1/n}$$

$$14. \sum_{n=1}^{\infty} (-1)^{n+1} \frac{3^n}{n2^n}$$

$$15. \sum_{n=0}^{\infty} (-1)^{n+1} \frac{3n-1}{4^n}$$

$$16. \sum_{n=0}^{\infty} (-1)^{n+1} \frac{3n^{2/3} - 1}{4^n}$$

$$17. \sum_{n=0}^{\infty} (-1)^{n+1} \frac{10}{(5n-2)^{1/4}}$$

$$18. \sum_{n=0}^{\infty} (-1)^{n+1} \frac{(n-3)!}{3^{n-1}}$$

$$19. \sum_{n=1}^{\infty} (-1)^{n+1} \frac{n+2}{n+1}$$

$$20. \sum_{n=1}^{\infty} \frac{n^3}{n!}$$

$$21. \sum_{n=1}^{\infty} \frac{3^{2n-1}}{n^2 + 1}$$

$$22. \sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{\sqrt{n}}$$

Answers to Odd-Numbered Problems

- | | |
|--------------------------|------------------------------|
| 1. absolutely convergent | 13. divergent |
| 3. divergent | 15. absolutely convergent |
| 5. absolutely convergent | 17. conditionally convergent |
| 7. absolutely convergent | 19. divergent |
| 9. divergent | 21. divergent |
| 11. divergent | |

POWER SERIES

An infinite series of the form

$$a_0 + a_1x + a_2x^2 + \cdots + a_nx^n + \cdots = \sum_{n=0}^{\infty} a_nx^n$$

where the coefficients a_0, a_1, a_2, \dots are independent of x , is called a *power series in x* . More generally, an infinite series of the form

$$b_0 + b_1(x-a) + b_2(x-a)^2 + \cdots + b_n(x-a)^n + \cdots = \sum_{n=0}^{\infty} b_n(x-a)^n$$

where the coefficients b_0, b_1, b_2, \dots are independent of x , is called a *power series in $x-a$* .

A power series in x (or in $x-a$) may converge for all values of x , or for no values of x other than $x=0$ (or $x=a$); or it may converge for some values of x and diverge for others. If a power series converges for values of x in the interval $-R < x < R$, then $-R < x < R$ is the *interval of convergence* of the power series and R is the *radius of convergence*. The interval of convergence is determined by the following procedure, derived from the Cauchy test-ratio test.

Power series in x :

$$\text{If } \lim_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} = L, \text{ then } \begin{cases} L = 0 \Rightarrow \text{series converges for all } x \\ L \neq 0 \Rightarrow \text{series converges for the interval} \\ \quad -1/|L| < x < 1/|L| \text{ and diverges outside this inter-} \\ \quad \text{val; the end points of the interval of convergence} \\ \quad \text{must be examined separately.} \end{cases}$$

Power series in $x-a$:

$$\text{If } \lim_{n \rightarrow \infty} \frac{b_{n+1}}{b_n} = M, \text{ then } \begin{cases} M = 0 \Rightarrow \text{series converges for all } x \\ M \neq 0 \Rightarrow \text{series converges for the interval} \\ \quad a - (1/|M|) < x < a + (1/|M|) \text{ and diverges out-} \\ \quad \text{side this interval; the end points of the interval of} \\ \quad \text{convergence must be examined separately.} \end{cases}$$

EXAMPLE

Find the interval of convergence for the power series

$$1 + 2x + 3x^2 + 4x^3 + \cdots$$

General term:

$$u_n = nx^{n-1}$$

$$L = \lim_{n \rightarrow \infty} \frac{n+1}{n} = 1,$$

so the interval of convergence is $-1 < x < 1$ and the end points must be tested.

If $x = -1$, the series is $1 - 2 + 3 - 4 + \cdots$

$$u_n = (-1)^{n-1}n$$

$$\lim_{n \rightarrow \infty} u_n \neq 0$$

so the series is divergent.

If $x = 1$, the series is $1 + 2 + 3 + 4 + \cdots$

$$u_n = n$$

$$\lim_{n \rightarrow \infty} u_n \neq 0$$

so the series is divergent.

Thus the original power series converges for $-1 < x < 1$.

EXAMPLE

Find the interval of convergence for the power series

$$1 - \frac{x}{2} + \frac{x^2}{3 \cdot 2^2} - \frac{x^3}{5 \cdot 2^3} + \frac{x^4}{7 \cdot 2^4} - \cdots$$

General term:

$$u_n = \frac{(-1)^{n+1}x^{n-1}}{(2n-3)2^{n-1}}$$

$$L = \lim_{n \rightarrow \infty} \left(-\frac{(2n-3)2^{n-1}}{[2(n+1)-3]2^n} \right) = \lim_{n \rightarrow \infty} \left(-\frac{2n-3}{4n-2} \right) = -\frac{1}{2}$$

so the interval of convergence is $-2 < x < 2$ and the end points must be tested.

If $x = -2$, the series is

$$1 + \frac{2}{2} + \frac{2^2}{3 \cdot 2^2} + \frac{2^3}{5 \cdot 2^3} + \frac{2^4}{7 \cdot 2^4} + \cdots$$

or

$$1 + 1 + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \cdots$$

$$u_n = \frac{1}{2n-1} \quad (\text{omitting the first term})$$

$$\lim_{n \rightarrow \infty} u_n = 0$$

$$\lim_{n \rightarrow \infty} \frac{u_{n+1}}{u_n} = \lim_{n \rightarrow \infty} \frac{\frac{1}{2(n+1)-1}}{\frac{1}{2n-1}} = \lim_{n \rightarrow \infty} \frac{2n-1}{2n+1} = 1$$

so the ratio test fails.

$$u_n = \frac{1}{2n-1} > \frac{1}{2n} \quad (p\text{-series, } p = 1)$$

so the series is divergent.

NOTE: Multiplying a series by a constant does *not* affect convergence.

If $x = 2$, the series is

$$1 - \frac{2}{2} + \frac{2^2}{3 \cdot 2^2} - \frac{2^3}{5 \cdot 2^3} + \frac{2^4}{7 \cdot 2^4} - \cdots$$

or

$$1 - 1 + \frac{1}{3} - \frac{1}{5} + \frac{1}{7} - \cdots$$

$$u_n = (-1)^n \frac{1}{2n-1} \quad (\text{omitting first term})$$

$$\lim_{n \rightarrow \infty} u_n = 0$$

$$|u_{n+1}| < |u_n|$$

so the alternating series is convergent.

Thus the original power series converges for $-2 < x \leq 2$.

EXAMPLE

Find the interval of convergence for the power series

$$1 + x + 2!x^2 + 3!x^3 + \cdots$$

General term:

$$u_n = (n-1)! x^{n-1}$$

$$L = \lim_{n \rightarrow \infty} \frac{n!}{(n-1)!} = \lim_{n \rightarrow \infty} n = \infty$$

Thus the power series converges only for $x = 0$, that is, the series is divergent for all $x \neq 0$.

EXAMPLE

Find the interval of convergence for the power series

$$(x-1) - \frac{1}{2}(x-1)^2 + \frac{1}{3}(x-1)^3 - \frac{1}{4}(x-1)^4 + \cdots$$

General term:

$$u_n = (-1)^{n+1} \frac{(x-1)^n}{n}$$

$$M = \lim_{n \rightarrow \infty} \left(-\frac{n}{n+1} \right) = -1$$

so the interval of convergence is $0 < x < 2$ and the end points must be tested.

If $x = 0$, the series is $-1 - \frac{1}{2} - \frac{1}{3} - \frac{1}{4} - \cdots$.

$$u_n = -\frac{1}{n}$$

$$\lim_{n \rightarrow \infty} u_n = 0$$

$$\lim_{n \rightarrow \infty} \frac{u_{n+1}}{u_n} = 1$$

so the ratio test fails.

$$u_n = \frac{1}{n} \quad (p\text{-series, } p = 1)$$

so the series is divergent.

If $x = 2$, the series is $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots$.

$$u_n = (-1)^{n+1} \frac{1}{n}$$

$$\lim_{n \rightarrow \infty} u_n = 0$$

$$|u_{n+1}| < |u_n|$$

so the alternating series is convergent.

Thus the original power series converges for $0 < x \leq 2$.

EXAMPLE

Find the interval of convergence for the power series

$$(x-1) - \frac{1}{4}(x-1)^2 + \frac{1}{9}(x-1)^3 - \frac{1}{16}(x-1)^4 + \cdots$$

General term:

$$u_n = (-1)^{n+1} \frac{(x-1)^n}{n^2}$$

$$M = \lim_{n \rightarrow \infty} \left(-\frac{n^2}{(n+1)^2} \right) = -1$$

so the interval of convergence is $0 < x < 2$ and the end points must be tested.

If $x = 0$, the series is $-1 - \frac{1}{4} - \frac{1}{9} - \frac{1}{16} - \cdots$.

$$u_n = -\frac{1}{n^2}$$

$$\lim_{n \rightarrow \infty} u_n = 0$$

$$\lim_{n \rightarrow \infty} \frac{u_{n+1}}{u_n} = 1$$

so the ratio test fails.

$$|u_n| = \frac{1}{n^2} \quad (p\text{-series, } p = 2)$$

so the series is convergent.

If $x = 2$, the series is $1 - \frac{1}{4} + \frac{1}{9} - \frac{1}{16} + \cdots$.

$$u_n = (-1)^{n+1} \frac{1}{n^2}$$

so the series is absolutely convergent (by procedure above).

Thus the original power series converges for $0 \leq x \leq 2$.

EXAMPLE

Find the interval of convergence for the power series

$$2(x-2) + \frac{3(x-2)^2}{2!} + \frac{4(x-2)^3}{3!} + \frac{5(x-2)^4}{4!} + \cdots$$

General term:

$$u_n = \frac{(n+1)(x-2)^n}{n!}$$

$$M = \lim_{n \rightarrow \infty} \left(\frac{\frac{n+2}{(n+1)!}}{\frac{n+1}{n!}} \right) = \lim_{n \rightarrow \infty} \frac{n+2}{(n+1)^2} = 0$$

Thus the power series converges for all x .

TAYLOR'S THEOREM

For either theoretical or computational purposes, it is frequently convenient to represent a function of x by a power series; Taylor's theorem provides appropriate power series for representing many functions.

TAYLOR'S THEOREM: The infinite series

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (x-a)^k = f(a) + f'(a) \frac{x-a}{1!} + f''(a) \frac{(x-a)^2}{2!} + \cdots + f^{(n-1)}(a) \frac{(x-a)^{n-1}}{(n-1)!} + R_n$$

converges and represents the function $f(x)$ for those values of x for which all the derivatives of $f(x)$ exist and for which $R_n \rightarrow 0$ as $n \rightarrow \infty$. In this case $f(x)$ is said to be expanded in a *Taylor's series about $x = a$* . For the special case $a = 0$, the expansion is a *Maclaurin's series*.

R_n is called the *remainder after n terms* and it can be shown that

$$R_n = f^{(n)}(\xi) \frac{(x-a)^n}{n!} \quad \text{where } a \leq \xi \leq x$$

This formula can be used to determine a bound for the error involved in using only the first n terms of the series if $\max_{a \leq \xi \leq x} R_n$ is obtained.

NOTE: There are some functions for which Taylor's series converges for values of x for which the remainder does not approach 0 as $n \rightarrow \infty$; for such values of x , the series does not represent the function. However, in most cases the interval of convergence of the series is the same as the interval for which $R_n \rightarrow 0$ as $n \rightarrow \infty$, as is the case for the examples considered here.

A Taylor's series about $x = a$ is useful for calculating the function it represents for values of x near a . Similarly, a Maclaurin series is useful for calculating the function it represents for values of x near zero.

A rigorous proof of Taylor's theorem is not simple and will not be given. However, the result can be made plausible by the argument presented in Technical Note III.

It should be noted that any differentiable function can be expanded in a Taylor's series: A function that can be differentiated only a finite number of times can be expanded in a series having a finite number of terms; a function that can be differentiated indefinitely can be expanded in an infinite series. However, to be a valid representation of a function the series must converge for values of x in the range of interest; in addition, to be useful for computational purposes, the series must converge fairly rapidly, so a reasonably accurate approximation can be obtained using a manageable number of terms. Thus, before using a Taylor's series expansion to represent a function, its convergence properties should be investigated.

EXAMPLE

Expand the function $f(x) = x^{-1/3}$ in powers of $(x - 1)$ and determine the interval of convergence of the series.

$$f(x) = x^{-1/3}$$

$$f'(x) = -\frac{1}{3}x^{-4/3}$$

$$f''(x) = \frac{1 \cdot 4}{3^2} x^{-7/3}$$

$$f'''(x) = -\frac{1 \cdot 4 \cdot 7}{3^3} x^{-10/3}$$

$$f^{(IV)}(x) = \frac{1 \cdot 4 \cdot 7 \cdot 10}{3^4} x^{-13/3}$$

$$f^{(V)}(x) = -\frac{1 \cdot 4 \cdot 7 \cdot 10 \cdot 13}{3^5} x^{-16/3}$$

$$\vdots$$

$$f^{(n)}(x) = (-1)^n \frac{(3n-2)(3n-5) \cdots 1}{3^n} x^{-(3n+1)/3}$$

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \cdots$$

$$= 1 - \frac{1}{3 \cdot 1!}(x-1) + \frac{1 \cdot 4}{3^2 \cdot 2!}(x-1)^2 - \frac{1 \cdot 4 \cdot 7}{3^3 \cdot 3!}(x-1)^3 + \cdots$$

$$+ (-1)^n \frac{(3n-2)(3n-5) \cdots 1}{3^n \cdot n!}(x-1)^n + \cdots$$

$$\frac{b_{n+1}}{b_n} = -\frac{\frac{(3n+1)(3n-2) \cdots 1}{3^{n+1}(n+1)!}}{\frac{(3n-2)(3n-5) \cdots 1}{3^n n!}} = -\frac{3n+1}{3(n+1)}$$

$\lim_{n \rightarrow \infty} b_{n+1}/b_n = -1$, so the interval of convergence is $0 < x < 2$ and the end points must be tested.

If $x = 0$,

$$u_n = \frac{(3n-2)(3n-5) \cdots 1}{3^n n!}$$

$$\lim_{n \rightarrow \infty} u_n = 1 \neq 0$$

so the series is divergent.

NOTE: For notational consistency, the general term is denoted u_n ; actually it is the term involving the n th derivative and is the $(n+1)$ st term of the series.

If $x = 2$,

$$u_n = (-1)^n \frac{(3n-2)(3n-5) \cdots 1}{3^n n!}$$

$$\lim_{n \rightarrow \infty} u_n = 1 \neq 0$$

so the series is divergent.

Thus the power series expansion converges for $0 < x < 2$.

EXAMPLE

Expand the function $f(x) = 1/(1+x)^2$ in powers of $(x-2)$ and determine the interval of convergence of the series.

$$f(x) = (1+x)^{-2}$$

$$f'(x) = -2(1+x)^{-3}$$

$$f''(x) = 2 \cdot 3(1+x)^{-4}$$

$$f'''(x) = -2 \cdot 3 \cdot 4(1+x)^{-5}$$

$$f^{(IV)}(x) = 2 \cdot 3 \cdot 4 \cdot 5(1+x)^{-6}$$

$$f^{(V)}(x) = -2 \cdot 3 \cdot 4 \cdot 5 \cdot 6(1+x)^{-7}$$

\vdots

$$f^{(n)}(x) = (-1)^n (n+1)! (1+x)^{-(n+2)}$$

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \cdots$$

$$= \frac{1}{3^2} - \frac{2}{3^3}(x-2) + \frac{3}{3^4}(x-2)^2 - \frac{4}{3^5}(x-2)^3 + \cdots + (-1)^n \frac{n+1}{3^{n+2}}(x-2)^n + \cdots$$

$$\frac{b_{n+1}}{b_n} = -\frac{\frac{n+2}{3^{n+3}}}{\frac{n+1}{3^{n+2}}} = -\frac{n+2}{3(n+1)}$$

$\lim_{n \rightarrow \infty} b_{n+1}/b_n = -\frac{1}{3}$ so the interval of convergence is $-1 < x < 5$ and the end points must be tested.

If $x = -1$, $u_n = (n+1)/3^2$.

$$\lim_{n \rightarrow \infty} u_n = \infty \neq 0,$$

so the series is divergent.

If $x = 5$, $u_n = (-1)^n (n+1)/3^2$.

$$\lim_{n \rightarrow \infty} u_n = \infty \neq 0,$$

so the series is divergent.

Thus the power series expansion converges for $-1 < x < 5$.

EXAMPLE

Expand the function $f(x) = (1+x)^m$, where m is any real number, in powers of x and determine the interval of convergence of the series. This is known as the *binomial expansion*.

$$f(x) = (1+x)^m$$

$$f'(x) = m(1+x)^{m-1}$$

$$f''(x) = m(m-1)(1+x)^{m-2}$$

$$f'''(x) = m(m-1)(m-2)(1+x)^{m-3}$$

$$f^{(IV)}(x) = m(m-1)(m-2)(m-3)(1+x)^{m-4}$$

$$f^{(V)}(x) = m(m-1)(m-2)(m-3)(m-4)(1+x)^{m-5}$$

$$\vdots$$

$$f^{(n)}(x) = \frac{m!}{(m-n)!} (1+x)^{m-n}$$

$$f(x) = f(a) + \frac{f'(a)}{1!}x + \frac{f''(a)}{2!}x^2 + \cdots + \frac{f^{(n)}(a)}{n!}x^n + \cdots$$

$$= 1 + \frac{m}{1!}x + \frac{m(m-1)}{2!}x^2 + \frac{m(m-1)(m-2)}{3!}x^3 + \cdots$$

$$+ \frac{m(m-1)(m-2) \cdots (m-n+1)}{n!}x^n + \cdots$$

$$\frac{b_{n+1}}{b_n} = \frac{\frac{m(m-1)(m-2) \cdots (m-n)}{(n+1)!}}{\frac{m(m-1)(m-2) \cdots (m-n+1)}{n!}} = \frac{m-n}{n+1}$$

$\lim_{n \rightarrow \infty} b_{n+1}/b_n = -1$, so the interval of convergence is $-1 < x < 1$ and the end points must be tested.

$$\text{If } x = -1, u_n = (-1)^n \frac{m(m-1)(m-2) \cdots (m-n+1)}{n!}.$$

$$\lim_{n \rightarrow \infty} u_n = 0$$

$$\text{If } x = 1, u_n = \frac{m(m-1)(m-2) \cdots (m-n+1)}{n!}.$$

$$\lim_{n \rightarrow \infty} u_n = 0$$

If $x = \pm 1$, $|u_{n+1}/u_n| = |(m-n)/(n+1)|$, so $|u_{n+1}| < |u_n|$ does not hold for unrestricted m .

Thus the power series expansion converges for $-1 < x < 1$.

EXAMPLE

Expand the function $f(x) = \ln x$ in a Taylor's series about $x = 1$ and determine the interval of convergence; use this series to evaluate $\ln 1.04$ correct to five decimal places.

$$f(x) = \ln x$$

$$f'(x) = \frac{1}{x}$$

$$f''(x) = -\frac{1}{x^2}$$

$$f'''(x) = \frac{2}{x^3}$$

$$f^{(IV)}(x) = -\frac{6}{x^4}$$

$$f^{(V)}(x) = \frac{24}{x^5}$$

$$\vdots$$

$$f^n(x) = (-1)^{n+1} \frac{(n-1)!}{x^n}$$

$$\ln x = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n-1)}(a)}{(n-1)!}(x-a)^{n-1} + R_n$$

$$= 0 + (x-1) - \frac{1}{2}(x-1)^2 + \frac{1}{3}(x-1)^3 - \cdots + (-1)^n \left(\frac{1}{n-1} \right) (x-1)^{n-1} + R_n$$

$$\frac{b_{n+1}}{b_n} = -\frac{\frac{1}{n}}{\frac{1}{n-1}} = -\frac{n-1}{n}$$

$\lim_{n \rightarrow \infty} b_{n+1}/b_n = -1$, so the interval of convergence is $0 < x < 2$ and the end points must be tested.

$$\text{If } x = 0, u_n = -\frac{1}{n-1}$$

$$\lim_{n \rightarrow \infty} u_n = 0$$

$$\lim_{n \rightarrow \infty} \frac{u_{n+1}}{u_n} = \lim_{n \rightarrow \infty} \frac{\frac{1}{n}}{\frac{1}{n-1}} = \lim_{n \rightarrow \infty} \frac{n-1}{n} = 1$$

so the ratio test fails. But

$$u_n = -\frac{1}{n-1} \quad (p\text{-series, } p = 1),$$

so the series is divergent.

NOTE: Omitting a finite number of terms from a series does not affect convergence, nor does multiplying the series by a positive or negative constant.

$$\text{If } x = 2, u_n = (-1)^{n+1}(1/n).$$

$$\lim_{n \rightarrow \infty} u_n = 0$$

$$|u_{n+1}| < |u_n|$$

so the series is conditionally convergent.

Thus the power series expansion converges for $0 < x \leq 2$.

$$\begin{aligned}
 \ln(1.04) &= 0 + 0.04 - \frac{1}{2}(0.04)^2 + \frac{1}{3}(0.04)^3 + R_4 \\
 &= 0.04 - 0.0008 + 0.000021 + R_4 \\
 &= 0.039221 + R_4
 \end{aligned}$$

$$R_n = f^n(\xi) \frac{(x-a)^n}{n!} \quad a \leq \xi \leq x$$

$$\begin{aligned}
 R_4 &= (-1)^5 \frac{3!}{\xi^4} \frac{(0.04)^4}{4!} \quad 1 \leq \xi \leq 1.04 \\
 &= (-1)^5 \frac{(0.04)^4}{4\xi^4}
 \end{aligned}$$

$$|R_4| \leq \frac{1}{4}(0.04)^4 = 0.00000064$$

so $\ln 1.04 = 0.03922$, correct to five decimal places.

NOTE: Since the series is alternating, the error in truncating is less in absolute value than the next term, which for this example gives the same bound on the error as above.

EXAMPLE

Expand the function $f(x) = e^x$ in a Maclaurin's series about $x = 0$ and determine the interval of convergence; use this series to evaluate $e^{1/2}$ correct to four decimal places.

$$f(x) = e^x$$

$$f'(x) = e^x$$

$$f''(x) = e^x$$

$$\vdots$$

$$f^n(x) = e^x$$

$$e^x = f(a) + \frac{f'(a)}{1!}x + \frac{f''(a)}{2!}x^2 + \cdots + \frac{f^{(n-1)}(a)}{(n-1)!}x^{n-1} + R_n$$

$$= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^{n-1}}{(n-1)!} + R_n$$

$$\frac{b_{n+1}}{b_n} = \frac{\frac{1}{n!}}{\frac{1}{(n-1)!}} = \frac{(n-1)!}{n!} = \frac{1}{n}$$

$$\lim_{n \rightarrow \infty} \frac{b_{n+1}}{b_n} = 0$$

Thus the power series expansion converges for all x .

$$e^{1/2} = 1 + \frac{1}{2} + \frac{\left(\frac{1}{2}\right)^2}{2!} + \frac{\left(\frac{1}{2}\right)^3}{3!} + \frac{\left(\frac{1}{2}\right)^4}{4!} + \frac{\left(\frac{1}{2}\right)^5}{5!} + R_6$$

$$= 1 + 0.5 + 0.125 + 0.02083333 + 0.00260417 + 0.000260417 + R_6$$

$$= 1.648697917 + R_6$$

$$R_n = f^n(\xi) \frac{(x-a)^n}{n!} \quad a \leq \xi \leq x$$

$$R_6 = e^\xi \frac{(\frac{1}{2})^6}{6!} \quad 0 \leq \xi \leq \frac{1}{2}$$

$$|R_6| \leq \frac{(\frac{1}{2})^6}{6!} e^{1/2}$$

and, since $e^{1/2} < 2$, $|R_6| < 0.000043$, so $e^{1/2} = 1.6487$, correct to four decimal places.

EXAMPLE

Expand the function $f(x) = (1+x)^{1/2}$ in a Maclaurin's series about $x = 0$ and determine the interval of convergence; use this series to evaluate $\sqrt{1.10}$ correct to six decimal places.

$$f(x) = (1+x)^{1/2}$$

$$f'(x) = \frac{1}{2}(1+x)^{-1/2}$$

$$f''(x) = -\frac{1}{2^2}(1+x)^{-3/2}$$

$$f'''(x) = \frac{1 \cdot 3}{2^3}(1+x)^{-5/2}$$

$$f^{(iv)}(x) = -\frac{1 \cdot 3 \cdot 5}{2^4}(1+x)^{-7/2}$$

$$f^{(v)}(x) = \frac{1 \cdot 3 \cdot 5 \cdot 7}{2^5}(1+x)^{-9/2}$$

$$\vdots$$

$$f^n(x) = (-1)^{n+1} \frac{(2n-3)(2n-5) \cdots 1}{2^n} (1+x)^{-(2n-1)/2}$$

$$(1+x)^{1/2} = f(a) + \frac{f'(a)}{1!}x - \frac{f''(a)}{2!}x^2 + \cdots + \frac{f^{(n-1)}(a)}{(n-1)!}x^{n-1} + R_n$$

$$= 1 + \frac{1}{2 \cdot 1!}x - \frac{1}{2^2 \cdot 2!}x^2 + \frac{3 \cdot 1}{2^2 \cdot 3!}x^3 - \cdots$$

$$+ (-1)^n \frac{(2n-5)(2n-7) \cdots 1}{2^{n-1}(n-1)!}x^{n-1} + R_n$$

$$\frac{b_{n+1}}{b_n} = -\frac{\frac{(2n-3)(2n-5)(2n-7) \cdots 1}{2^n n!}}{\frac{(2n-5)(2n-7) \cdots 1}{2^{n-1}(n-1)!}} = -\frac{2n-3}{2n}$$

$\lim_{n \rightarrow \infty} b_{n+1}/b_n = -1$, so the interval of convergence is $-1 < x < 1$ and the end points must be tested.

If $x = -1$, $f(x) = 0$, so expansion is unnecessary.

$$\text{If } x = 1, u_n = (-1)^n \frac{(2n-5)(2n-7) \cdots 1}{2^{n-1}(n-1)!}.$$

$$\lim_{n \rightarrow \infty} u_n = 0$$

$$\left| \frac{u_{n+1}}{u_n} \right| = \frac{2n-3}{2n} < 1$$

so the alternating series is conditionally convergent.

Thus the power series expansion converges for $-1 < x \leq 1$.

$$\begin{aligned}(1.10)^{1/2} &= 1 + \frac{1}{2 \cdot 1!} \left(\frac{1}{10}\right) - \frac{1}{2^2 \cdot 2!} \left(\frac{1}{10}\right)^2 + \frac{3 \cdot 1}{2^3 \cdot 3!} \left(\frac{1}{10}\right)^3 - \frac{5 \cdot 3 \cdot 1}{2^4 \cdot 4!} \left(\frac{1}{10}\right)^4 \\ &\quad + \frac{7 \cdot 5 \cdot 3 \cdot 1}{2^5 \cdot 5!} \left(\frac{1}{10}\right)^5 + R_6 \\ &= 1 + 0.05 - 0.00125 + 0.0000625 - 0.000003906 + 0.000000273 + R_6 \\ &= 1.048808867 + R_6\end{aligned}$$

Alternating series, so the error is less than the absolute value of next term,

$$\frac{9 \cdot 7 \cdot 5 \cdot 3 \cdot 1}{2^6 \cdot 6!} \left(\frac{1}{10}\right)^6 = 0.0000000205$$

so correct to six decimal places, $(1.10)^{1/2} = 1.048809$.

NOTE: Alternatively, a bound on the error can be obtained using R_6 :

$$R_n = f^n(\xi) \frac{(x-a)^n}{n!} \quad a \leq \xi \leq x$$

$$R_6 = -\frac{9 \cdot 7 \cdot 5 \cdot 3 \cdot 1}{2^6 \cdot 6!} (1 + \xi)^{-11/2} \left(\frac{1}{10}\right)^6 \quad 0 \leq \xi \leq \frac{1}{10}$$

$$|R_6| \leq \frac{9 \cdot 7 \cdot 5 \cdot 3 \cdot 1}{2^6 \cdot 6!} \left(\frac{1}{10}\right)^6$$

$$|R_6| \leq 0.0000000205$$

Note that the accuracy of the approximation obtained from a power series expansion increases with the number of terms used in the calculation. Fewer terms are required for a specified accuracy if the series converges more rapidly. The number of terms required for a given accuracy must be determined for each power series approximation by considering the remainder term or, for an alternating series, by considering the next term. Note that the remainder after n terms

$$R_n = f^{(n)}(\xi) \frac{(x-a)^n}{n!} \quad a \leq \xi \leq x$$

is equal in absolute value to the next term of an alternating series if $\xi = a$. Thus, for an alternating series, the bound on the error determined by obtaining $\max_{a \leq \xi \leq x} R_n$ is equal to the bound determined by obtaining the absolute value of the next term of the alternating series, if $\max_{a \leq \xi \leq x} R_n$ occurs for $\xi = a$.

PROBLEMS

For each of the following power series, determine the general term and the interval of convergence.

$$1. x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$$

$$2. 2x + 4x^2 + 8x^3 + 16x^4 + \cdots$$

$$3. 1 + \frac{x}{\sqrt{2}} + \frac{x^2}{\sqrt{3}} + \frac{x^3}{\sqrt{4}} + \cdots$$

$$4. \frac{x}{1 \cdot 2} + \frac{x^2}{2^2 \cdot 2^2} + \frac{x^3}{3^2 \cdot 2^3} + \frac{x^4}{4^2 \cdot 2^4} + \cdots$$

$$5. 1 - \frac{3x}{2} + \frac{5x^2}{4} - \frac{7x^3}{8} + \frac{9x^4}{16} - \dots$$

$$6. 1 + \frac{(x+3)}{2^2} + \frac{(x+3)^2}{3^2} + \frac{(x+3)^3}{4^2} + \dots$$

$$7. \frac{1}{2!}x + \frac{2}{3!}x^3 + \frac{3}{4!}x^5 + \frac{4}{5!}x^7 + \dots$$

$$8. \frac{x+4}{\sqrt[3]{2}} - \frac{(x+4)^2}{\sqrt[3]{3}} + \frac{(x+4)^3}{\sqrt[3]{4}} - \frac{(x+4)^4}{\sqrt[3]{5}} + \dots$$

$$9. \frac{1}{2}x - \frac{2^2}{5}x^2 + \frac{3^2}{8}x^3 - \frac{4^2}{11}x^4 + \frac{5^2}{14}x^5 + \dots$$

$$10. \frac{6}{2}(x-5) + \frac{6^2}{5}(x-5)^2 + \frac{6^3}{10}(x-5)^3 + \frac{6^4}{17}(x-5)^4 + \frac{6^5}{26}(x-5)^5 + \dots$$

Expand the following functions in power series, determine the interval of convergence, and choose some appropriate value of x for which to estimate the function.

$$11. f(x) = x^{1/2} \text{ in powers of } (x-1)$$

$$17. f(x) = \sin x \text{ in powers of } x$$

$$12. f(x) = \frac{1}{1+x^2} \text{ in powers of } x$$

$$18. f(x) = \cos x \text{ in powers of } x$$

$$13. f(x) = e^{-(1/2)x} \text{ in powers of } (x-2)$$

$$19. f(x) = \tan x \text{ in powers of } x$$

$$14. f(x) = \frac{1}{2}(e^x + e^{-x}) \text{ in powers of } x$$

$$20. f(x) = \cot x \text{ in powers of } x$$

$$15. f(x) = \ln(a-x) \text{ in powers of } x$$

$$21. f(x) = \sec x \text{ in powers of } x$$

$$16. f(x) = \ln(a+x) \text{ in powers of } x$$

$$22. f(x) = \csc x \text{ in powers of } x$$

Answers to Odd-Numbered Problems

$$1. (-1)^{n+1} \frac{x^{2n-1}}{(2n-1)!}, \text{ conv. all } x$$

$$3. \frac{x^{n-1}}{\sqrt{n}}, \text{ conv. } -1 \leq x < 1$$

$$5. (-1)^{n-1} \frac{(2n-1)x^{n-1}}{2^{n-1}}, \text{ conv. } -2 < x < 2$$

$$7. \frac{n}{(n+1)!} x^{2n-1}, \text{ conv. all } x$$

$$9. (-1)^{n+1} \frac{n^2}{3n-1} x^n, \text{ conv. } -1 < x < 1$$

$$\begin{aligned} 11. f(x) = 1 + \frac{1}{2 \cdot 1!} (x-1) - \frac{1}{2^2 \cdot 2!} (x-1)^2 + \frac{1 \cdot 3}{2^3 \cdot 3!} (x-1)^3 \\ - \frac{1 \cdot 3 \cdot 5}{2^4 \cdot 4!} (x-1)^4 + \frac{1 \cdot 3 \cdot 5 \cdot 7}{2^5 \cdot 5!} (x-1)^5 + \dots \\ + (-1)^{n+1} \frac{(2n-3)(2n-5) \cdots 1}{2^n \cdot n!} (x-1)^n + \dots, \quad \text{conv. } 0 < x \leq 2 \end{aligned}$$

$$\begin{aligned} 13. f(x) = 1 - \frac{e^{-1}}{2 \cdot 1!} (x-2) + \frac{e^{-1}}{2^2 \cdot 2!} (x-2)^2 - \frac{e^{-1}}{2^3 \cdot 3!} (x-2)^3 \\ + \frac{e^{-1}}{2^4 \cdot 4!} (x-2)^4 - \frac{e^{-1}}{2^5 \cdot 5!} (x-2)^5 + \dots \\ + (-1)^n \frac{e^{-1}}{2^n \cdot n!} (x-2)^n + \dots, \text{ conv. all } x \end{aligned}$$

$$15. f(x) = \ln a - \frac{1}{a}x - \frac{1}{2a^2}x^2 - \frac{1}{3a^3}x^3 - \frac{1}{4a^4}x^4 - \cdots - \frac{1}{na^n}x^n - \cdots, \quad \text{conv. } -a \leq x < a$$

$$17. f(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots + (-1)^{n+1} \frac{x^{2n-1}}{(2n-1)!} + \cdots, \quad \text{conv. all } x$$

$$19. f(x) = x + \frac{x^3}{3} + \frac{2x^5}{15} + \frac{17x^7}{315} + \cdots, \quad \text{conv. } |x| < \frac{\pi}{2}$$

$$21. f(x) = 1 + \frac{x^2}{2} + \frac{5x^4}{24} + \frac{61x^6}{720} + \cdots, \quad \text{conv. } |x| < \frac{\pi}{2}$$

Technical Note I Derivation of Rules for Differentiation

NOTE: For mathematical convenience, the rules for differentiation are not derived in the same order as they are given in the text.

Algebraic Functions

Rule 1. If $y = c$,

$$\Delta y = 0$$

$$\frac{\Delta y}{\Delta x} = 0$$

$$\frac{dy}{dx} = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = 0$$

Rule 3. If $y = cu$, where $u = f(x)$ is a differentiable function of x ,

$$y + \Delta y = c(u + \Delta u)$$

$$\Delta y = c \Delta u$$

$$\frac{\Delta y}{\Delta x} = c \frac{\Delta u}{\Delta x}$$

$$\frac{dy}{dx} = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = c \frac{du}{dx}$$

Rule 2. If $y = x^n$, where n is a positive integer,

$$y + \Delta y = (x + \Delta x)^n$$

$$= x^n + nx^{n-1} \Delta x + \frac{n(n-1)}{2!} x^{n-2} (\Delta x)^2 + \cdots + (\Delta x)^n$$

$$\Delta y = nx^{n-1} \Delta x + \frac{n(n-1)}{2!} x^{n-2} (\Delta x)^2 + \cdots + (\Delta x)^n$$

$$\frac{\Delta y}{\Delta x} = nx^{n-1} + \frac{n(n-1)}{2!} x^{n-2} (\Delta x) + \cdots + (\Delta x)^{n-1}$$

$$\frac{dy}{dx} = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = nx^{n-1}$$

Rule 4. If $y = u + v$, where $u = f(x)$ and $v = g(x)$ are differentiable functions of x ,

$$y + \Delta y = u + \Delta u + v + \Delta v$$

$$\Delta y = \Delta u + \Delta v$$

$$\frac{\Delta y}{\Delta x} = \frac{\Delta u}{\Delta x} + \frac{\Delta v}{\Delta x}$$

$$\frac{dy}{dx} = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \frac{du}{dx} + \frac{dv}{dx}$$

Rule 5. If $y = uv$, where $u = f(x)$ and $v = g(x)$ are differentiable functions of x ,

$$\begin{aligned} y + \Delta y &= (u + \Delta u)(v + \Delta v) \\ &= uv + u(\Delta v) + v(\Delta u) + (\Delta u)(\Delta v) \\ \Delta y &= u(\Delta v) + v(\Delta u) + (\Delta u)(\Delta v) \\ \frac{\Delta y}{\Delta x} &= u \frac{\Delta v}{\Delta x} + v \frac{\Delta u}{\Delta x} + \frac{(\Delta u)(\Delta v)}{\Delta x} \\ \frac{dy}{dx} &= \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = u \frac{dv}{dx} + v \frac{du}{dx} \end{aligned}$$

Rule 6. If $y = u/v$, where $u = f(x)$ and $v = g(x)$ are differentiable functions of x ,

$$\begin{aligned} y + \Delta y &= \frac{u + \Delta u}{v + \Delta v} \\ \Delta y &= \frac{v(u + \Delta u) - u(v + \Delta v)}{v(v + \Delta v)} \\ &= \frac{v(\Delta u) - u(\Delta v)}{v(v + \Delta v)} \\ \frac{\Delta y}{\Delta x} &= \frac{v \frac{\Delta u}{\Delta x} - u \frac{\Delta v}{\Delta x}}{v(v + \Delta v)} \\ \frac{dy}{dx} &= \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \frac{v \frac{du}{dx} - u \frac{dv}{dx}}{v^2} \end{aligned}$$

Logarithmic Functions

Rule 8. If $y = \log_a u$, where $u = f(x)$ is a differentiable function of x ,

$$\begin{aligned} y + \Delta y &= \log_a (u + \Delta u) \\ \Delta y &= \log_a (u + \Delta u) - \log_a u \\ &= \log_a \left(\frac{u + \Delta u}{u} \right) \\ &= \log_a \left(1 + \frac{\Delta u}{u} \right) \\ \frac{\Delta y}{\Delta u} &= \frac{1}{\Delta u} \log_a \left(1 + \frac{\Delta u}{u} \right) \\ &= \frac{1}{u} \cdot \frac{u}{\Delta u} \log_a \left(1 + \frac{\Delta u}{u} \right) \\ &= \frac{1}{u} \log_a \left(1 + \frac{\Delta u}{u} \right)^{u/\Delta u} \end{aligned}$$

Let $\Delta u/u = h$; then $u/\Delta u = 1/h$ and $h \rightarrow 0$ as $\Delta u \rightarrow 0$. Thus $\Delta y/\Delta u = (1/u) \log_a (1 + h)^{1/h}$ and, since (by definition) $(1 + h)^{1/h} \rightarrow e$ as $h \rightarrow 0$,

$$\begin{aligned} \lim_{\Delta u \rightarrow 0} \frac{\Delta y}{\Delta u} &= \frac{1}{u} \log_a \left[\lim_{h \rightarrow 0} (1 + h)^{1/h} \right] \\ \frac{dy}{du} &= \frac{1}{u} \log_a e \end{aligned}$$

If $u = f(x)$,

$$\frac{dy}{dx} = \frac{1}{u} \log_a e \frac{du}{dx}$$

Exponential Functions

Rule 9. If $y = a^u$, where $u = f(x)$ is a differentiable function of x ,

$$\ln y = \ln a^u$$

$$= u \ln a$$

$$\frac{1}{y} \frac{dy}{dx} = \frac{du}{dx} \ln a$$

$$\frac{dy}{dx} = y \frac{du}{dx} \ln a$$

$$= a^u \ln a \frac{du}{dx}$$

Rule 10. If $y = u^v$, where $u = f(x)$ and $v = g(x)$ are differentiable functions of x ,

$$\ln y = v \ln u$$

$$\frac{1}{y} \frac{dy}{dx} = v \frac{1}{u} \frac{du}{dx} + \ln u \frac{dv}{dx}$$

$$\frac{dy}{dx} = vu^{v-1} \frac{du}{dx} + u^v \ln u \frac{dv}{dx}$$

Rule 7. If $y = u^n$, where $u = f(x)$ is a differentiable function of x and n is any real number,

$$\ln y = n \ln u$$

$$\frac{1}{y} \frac{dy}{dx} = n \frac{1}{u} \frac{du}{dx}$$

$$\frac{dy}{dx} = nu^{n-1} \frac{du}{dx}$$

Trigonometric Functions

Rule 11. If $y = \sin u$, where $u = f(x)$ is a differentiable function of x ,

$$y + \Delta y = \sin (u + \Delta u)$$

$$y = \sin (u + \Delta u) - \sin u$$

$$= 2 \cos \left(u + \frac{\Delta u}{2} \right) \sin \frac{\Delta u}{2}$$

$$\text{since } \sin x - \sin y = 2 \cos \frac{1}{2}(x + y) \sin \frac{1}{2}(x - y)$$

$$\frac{\Delta y}{\Delta u} = 2 \cos \left(u + \frac{\Delta u}{2} \right) \frac{\sin (\Delta u/2)}{\Delta u}$$

$$= \cos \left(u + \frac{\Delta u}{2} \right) \frac{\sin (\Delta u/2)}{\Delta u/2}$$

$$\begin{aligned}\frac{dy}{du} &= \lim_{\Delta u \rightarrow 0} \frac{\Delta y}{\Delta u} = \lim_{\Delta u \rightarrow 0} \cos \left(u + \frac{\Delta u}{2} \right) \frac{\sin (\Delta u/2)}{\Delta u/2} \\ &= \cos u \text{ since } \lim_{\Delta u \rightarrow 0} \frac{\sin (\Delta u/2)}{\Delta u/2} = 1\end{aligned}$$

If $u = f(x)$,

$$\frac{dy}{dx} = \cos u \frac{du}{dx}$$

Rule 12. If $y = \cos u$, where $u = f(x)$ is a differentiable function of x ,

$$y = \sin \left(\frac{\pi}{2} - u \right)$$

and

$$\begin{aligned}\frac{dy}{dx} &= \cos \left(\frac{\pi}{2} - u \right) \left(-\frac{du}{dx} \right) \\ &= -\sin u \frac{du}{dx}\end{aligned}$$

Rule 13. If $y = \tan u$, where $u = f(x)$

$$y = \frac{\sin u}{\cos u}$$

and

$$\frac{dy}{dx} = \frac{\cos^2 u \frac{du}{dx} + \sin^2 u \frac{du}{dx}}{\cos^2 u}$$

but $\cos^2 u + \sin^2 u = 1$, so

$$\frac{dy}{dx} = \frac{\frac{du}{dx}}{\cos^2 u} = \sec^2 u \frac{du}{dx}$$

Inverse Functions

Rule 14. If $y = f(x)$ and $x = g(y)$ are inverse differentiable functions,

$$\begin{aligned}\frac{g(y + \Delta y) - g(y)}{\Delta y} &= \frac{h}{\Delta y} \quad \begin{array}{l} \text{where } x = g(y) \\ x + h = g(y + \Delta y) \end{array} \\ y + \Delta y &= f(x + h) \quad \text{since } f[g(x)] = x\end{aligned}$$

and thus

$$\begin{aligned}\Delta y &= f(x + h) - f(x) \\ \frac{g(y + \Delta y) - g(y)}{\Delta y} &= \frac{\Delta x}{f(x + h) - f(x)}\end{aligned}$$

As $\Delta y \rightarrow 0$, $\Delta x \rightarrow 0$ and

$$\lim_{\Delta y \rightarrow 0} \frac{g(y + \Delta y) - g(y)}{\Delta y} = g'(y) = \lim_{\Delta x \rightarrow 0} \frac{\Delta x}{f(x + h) - f(x)} = \frac{1}{f'(x)}$$

That is,

$$\frac{dx}{dy} = \frac{1}{\frac{dy}{dx}}$$

Composite Functions

Rule 15. If $y = \phi(u)$ and $u = F(x)$

$$\begin{aligned}\frac{\Delta y}{\Delta x} &= \frac{\Delta y}{\Delta u} \cdot \frac{\Delta u}{\Delta x} \\ \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} &= \frac{dy}{du} \cdot \frac{du}{dx}\end{aligned}$$

Technical Note II Conditions for a Cubic Cost Function

$b^2 - 3ac \leq 0$ is a sufficient but not a necessary condition for

$$y = ax^3 + bx^2 + cx + d$$

to have no relative maximum or minimum in the first quadrant. If $b^2 - 3ac > 0$, then

$$\frac{dy}{dx} = 0 \text{ if } x = \frac{-b \pm \sqrt{b^2 - 3ac}}{3a}.$$

$$x = \frac{-b \pm \sqrt{b^2 - 3ac}}{3a} > 0 \quad \text{since } a > 0, b < 0 \quad \text{and} \quad |b^2 - 3ac| < |b|$$

(note that $b^2 - 3ac > 0 \Rightarrow b \neq 0$).

$$y = a\left(\frac{-b \pm \sqrt{b^2 - 3ac}}{3a}\right)^3 + b\left(\frac{-b \pm \sqrt{b^2 - 3ac}}{3a}\right)^2 + c\left(\frac{-b \pm \sqrt{b^2 - 3ac}}{3a}\right) + d$$

Since $a > 0, b < 0, c \geq 0, d \geq 0$ and $(-b \pm \sqrt{b^2 - 3ac})/3a > 0$, the first term in the expression for y is positive, the second term is negative, and the third and fourth terms are positive or zero. Thus y may be positive or negative depending on the relative value of the coefficients a, b, c , and d , and (x, y) may lie in either the first or second quadrant. Since

$$\left. \frac{d^2 y}{dx^2} \right|_{x = \frac{-b \pm \sqrt{b^2 - 3ac}}{3a}} = 2(-b \pm \sqrt{b^2 - 3ac} + b) = \pm 2\sqrt{b^2 - 3ac}$$

and $b^2 - 3ac > 0$,

$$\left. \frac{d^2 y}{dx^2} \right|_{x = \frac{-b \pm \sqrt{b^2 - 3ac}}{3a}} \neq 0$$

and $y = ax^3 + bx^2 + cx + d$ has a relative maximum or minimum at the point for which $\frac{dy}{dx} = 0$. Thus, if $b^2 - 3ac > 0$, $y = ax^3 + bx^2 + cx + d$ has a relative maximum or minimum in the first quadrant only for certain values of the coefficients. For simplicity, only equations for which $b^2 - 3ac \leq 0$ are considered in the text.

Technical Note III Taylor's Theorem

Assume that the function $f(x)$ has a power series expansion and that the power series can be differentiated term by term.

$$f(x) = a_0 + a_1(x - a) + a_2(x - a)^2 + \cdots + a_n(x - a)^n + \cdots$$

If $x = a$, $f(a) = a_0$.

$$f'(x) = a_1 + 2a_2(x-a) + \cdots + na_n(x-a)^{n-1} + \cdots$$

If $x = a$, $f'(a) = a_1$.

$$f''(x) = 2a_2 + 3 \cdot 2a_3(x-a) + \cdots + n(n-1)a_n(x-a)^{n-2} + \cdots$$

If $x = a$, $f''(a) = 2a_2$.

$$f'''(x) = 3 \cdot 2a_3 + 4 \cdot 3 \cdot 2a_4(x-a) + \cdots + n(n-1)(n-2)a_n(x-a)^{n-3} + \cdots$$

If $x = a$, $f'''(a) = 3 \cdot 2a_3$.

$$\vdots$$

$$f^k(x) = k(k-1)a_k + (k+1)(k)(k-1)a_{k+1}(x-a) + \cdots$$

$$+ n(n-1) \cdots (n-k+1)a_n(x-a)^{n-k} + \cdots$$

If $x = a$, $f^k(a) = k! a_k$.

Thus

$$a_n = \frac{f^n(a)}{n!} \quad \text{for } n = 1, 2, \dots; a_0 = f(a)$$

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^n(a)}{n!}(x-a)^n + \cdots$$

Differential Calculus: Functions of More Than One Variable

3

3.1 Introduction

The discussion and examples in previous sections were concerned with functions of one variable, that is, functions of the explicit form $y = f(x)$ or the implicit form $f(x, y) = 0$. Such functions express a relationship between two variables, x and y , and implicitly assume that the processes being studied can be represented adequately in terms of only these two variables. Although in many cases this type of representation provides a reasonably accurate first approximation to reality, there are clearly many cases in which such a representation is so inadequate as to be virtually useless. In these cases it is necessary to express a relationship in terms of several variables or to express one variable as a function of more than one other variable. For example, in economic theory supply and demand of a commodity frequently depend not only on its price but also on the prices of related commodities, on income level, on time, and on various other factors. Profit depends not just on the output of one commodity, but on the output levels for several commodities and on the combination of several inputs.

Relations and functions of several variables can be defined by appropriate extensions of the corresponding definitions for two variables. Geometric representation is not possible for functions involving more than three variables, but, with this exception, the following discussion and definitions are appropriate for any finite number of variables. Partial differentiation is defined and several applications of partial derivatives in business and economics are discussed. Constrained and unconstrained maxima and minima of functions of two variables are considered and constrained utility maximization is discussed as an application of constrained maximization and minimization in business and economics.

FUNCTIONS OF MORE THAN ONE VARIABLE

Just as a point in two-dimensional space is represented by an ordered pair of real numbers, a point in three-dimensional space is represented by an ordered triple of real numbers and a point in n -dimensional space is represented by an ordered n -tuple of real numbers (x_1, x_2, \dots, x_n) . Thus a point in n -dimensional space is an ordered arrangement of n real numbers. The set of all points in n -dimensional space is

denoted by U_n . For convenience, a point in n -dimensional space is sometimes denoted by a single letter, say X .

A set of ordered pairs $\{(X, z)\}$, where $X \in U_n$, is a relation. If $(X_1, z_1) \in F$ and $(X_1, z_2) \in F$ implies that $z_1 = z_2$, then $\{(X, z)\}$ is a function of n variables. The domain of the function is the set of points $\{X\}$ in the space U_n and the range is the set of values $\{z\}$ appearing in the set of pairs in F . Note that the set of ordered pairs is a relation and, if to each point of the domain there corresponds only one point of the range, the relation is a function.

When $X \in U_2$, the pair (X, z) is an ordered triple, say (x, y, z) , which may be interpreted as a point in three-dimensional space. Functions whose domain is U_k , where $k > 2$, cannot be represented geometrically.

A function of several variables is frequently described by an equation expressing the correspondence between the coordinates of the point X and the functional value z . For this purpose, the functional notation $y = f(x)$ can be extended for functions defined on U_n . For example, if $X \in U_n$, $z = F(X)$ indicates that F is a function of n variables. If the domain is in U_2 ,

$$z = F(x, y) \quad \text{or} \quad z = F(X)$$

and z is a function of two variables. In general, if $X \in U_n$,

$$z = F(x_1, x_2, \dots, x_n) \quad \text{or} \quad z = F(X)$$

and z is a function of n variables. The coordinates of X are called *independent variables* and z is called a *dependent variable*. Implicit functions of several variables can be written, for example, as $f(x, y, z) = 0$ or, more generally, as $f(x_1, x_2, \dots, x_n, z) = 0$.

The definition of a function is extended to the case of n independent variables in the above discussion. Most of the definitions and formulas given subsequently for functions involving more than two variables are stated explicitly for functions involving three or, in some cases, four variables. This is for convenience. These definitions and formulas can also be extended to the general case of n variables, but additional notation is required to write the corresponding statements concisely. Several results concerning maxima and minima of functions of n variables, which are not obtained easily as generalizations of the corresponding results for two variables, are stated in matrix notation, after that notation is defined.

Continuity

A function $f(x, y)$ is said to be *continuous* at $x = a$, $y = b$, if the following three conditions are satisfied:

1. $f(a, b)$ exists.
2. $\lim_{x \rightarrow a, y \rightarrow b} f(x, y)$ exists.
3. $\lim_{x \rightarrow a, y \rightarrow b} f(x, y) = f(a, b)$, no matter in what manner $x \rightarrow a$, $y \rightarrow b$.

A function of x and y is said to be *continuous in a region* of the xy -plane if it is continuous at every point of the region. Analogous definitions are appropriate for functions of n variables.

3.2 Partial Differentiation

In this section partial differentiation, the total differential, and the total derivative are defined and illustrated. Partial differentiation of implicit functions is also considered and illustrated.

Consider the function z of two independent variables x and y ,

$$z = f(x, y)$$

If y is held constant, z is a function only of x and the derivative of z with respect to x can be computed. The derivative obtained in this way is the *partial derivative* of z with respect to x and is denoted by

$$\frac{\partial z}{\partial x} = \frac{\partial f}{\partial x} = \frac{\partial}{\partial x} f(x, y) = f_x(x, y) = f_x = z_x$$

Similarly, if x is held constant the partial derivative with respect to y can be computed and is denoted by

$$\frac{\partial z}{\partial y} = \frac{\partial f}{\partial y} = \frac{\partial}{\partial y} f(x, y) = f_y(x, y) = f_y = z_y$$

The partial derivative of z with respect to x is defined as

$$\frac{\partial z}{\partial x} = \lim_{\Delta x \rightarrow 0} \frac{\Delta z}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x}$$

and the partial derivative of z with respect to y is defined as

$$\frac{\partial z}{\partial y} = \lim_{\Delta y \rightarrow 0} \frac{\Delta z}{\Delta y} = \lim_{\Delta y \rightarrow 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y}$$

NOTE: In general, a function of any number of variables may have a partial derivative with respect to each of its variables.

EXAMPLE

If $z = 2x^2 + 3xy - 6y^2$,

$$\frac{\partial z}{\partial x} = 4x + 3y$$

$$\frac{\partial z}{\partial y} = 3x - 12y$$

EXAMPLE

If $z = xy + \ln x$,

$$\frac{\partial z}{\partial x} = y + \frac{1}{x}$$

$$\frac{\partial z}{\partial y} = x$$

EXAMPLE

If $z = (x + y) \sin (x - y)$,

$$\frac{\partial z}{\partial x} = \sin (x - y) + (x + y) \cos (x - y)$$

$$\frac{\partial z}{\partial y} = \sin (x - y) - (x + y) \cos (x - y)$$

EXAMPLE

If $z = \frac{x^3 - y^3}{xy}$, show that $x \frac{\partial z}{\partial x} + y \frac{\partial z}{\partial y} = z$.

$$z = x^2 y^{-1} - x^{-1} y^2$$

$$\frac{\partial z}{\partial x} = 2xy^{-1} + x^{-2}y^2$$

$$\frac{\partial z}{\partial y} = -x^2y^{-2} - 2x^{-1}y$$

$$x \frac{\partial z}{\partial x} + y \frac{\partial z}{\partial y} = 2x^2y^{-1} + x^{-1}y^2 - x^2y^{-1} - 2x^{-1}y^2$$

$$= x^2y^{-1} - x^{-1}y^2 = z$$

Just as a function of one variable can be represented by a curve in a plane, a function of two variables can be represented by a surface in space. Consider the surface represented by $z = f(x, y)$ in Figure 3.1. If a plane is passed through any point of P of the surface parallel to the xz -plane, it cuts the surface in the curve APB , along which y remains constant; the slope of the tangent line to APB at P represents the rate at which z changes with respect to x at the point P . Similar statements hold for a plane parallel to the yz -plane. Thus the partial derivatives have the following geometrical significance:

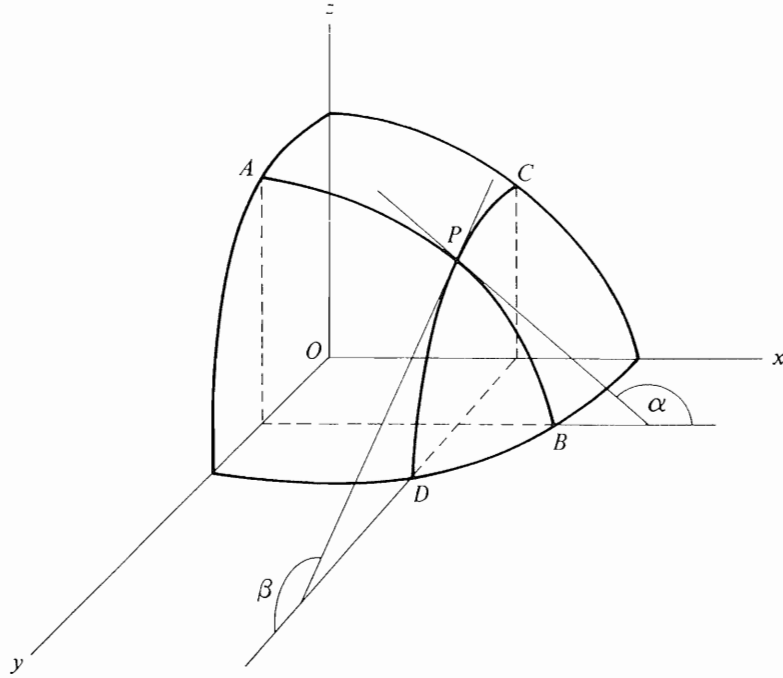
$$\left. \frac{\partial z}{\partial x} = \frac{\partial z}{\partial x} \right|_P = \tan \alpha = \text{slope of } APB \text{ at } P$$

$$\left. \frac{\partial z}{\partial y} = \frac{\partial z}{\partial y} \right|_P = \tan \beta = \text{slope of } CPD \text{ at } P$$

That is, if $z = f(x, y)$, the partial derivatives $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ are interpreted geometrically as the slopes of the curves of intersection of the surface $z = f(x, y)$ and the planes $y = \text{constant}$ and $x = \text{constant}$, respectively.

Since, in general, the partial derivatives of a function $z = f(x, y)$ are functions of x

Figure 3.1



and y , they may be differentiated with respect to either x or y ; these derivatives, if they exist, are called the *second partial derivatives* of z and are denoted, respectively, by

$$\frac{\partial}{\partial x} \left(\frac{\partial z}{\partial x} \right) = \frac{\partial^2 z}{\partial x^2} = z_{xx} = \frac{\partial^2 f}{\partial x^2} = f_{xx}$$

$$\frac{\partial}{\partial x} \left(\frac{\partial z}{\partial y} \right) = \frac{\partial^2 z}{\partial x \partial y} = z_{xy} = \frac{\partial^2 f}{\partial x \partial y} = f_{xy}$$

$$\frac{\partial}{\partial y} \left(\frac{\partial z}{\partial x} \right) = \frac{\partial^2 z}{\partial y \partial x} = z_{yx} = \frac{\partial^2 f}{\partial y \partial x} = f_{yx}$$

$$\frac{\partial}{\partial y} \left(\frac{\partial z}{\partial y} \right) = \frac{\partial^2 z}{\partial y^2} = z_{yy} = \frac{\partial^2 f}{\partial y^2} = f_{yy}$$

A second- or higher-order partial derivative which is obtained by differentiating with respect to more than one variable is referred to as a *mixed partial derivative*. Thus f_{xy} and f_{yx} are mixed partial derivatives. Of the four derivatives above, only three are generally distinct, since $f_{xy} = f_{yx}$ for all values of x and y for which f_{xy} and f_{yx} are continuous. Because of this property, it follows that a mixed partial derivative of second (or higher) order can be obtained by differentiating with respect to the variables in any order, provided only that the derivatives are continuous.

Similar definitions and equalities hold for higher-order partial derivatives of functions of two variables and for partial derivatives of functions of more than two variables.

EXAMPLE

If $z = (x^2 + y^2)^{3/2}$,

$$\frac{\partial z}{\partial x} = \frac{3}{2}(x^2 + y^2)^{1/2}(2x)$$

$$= 3x(x^2 + y^2)^{1/2}$$

$$\frac{\partial^2 z}{\partial y \partial x} = \frac{3}{2}x(x^2 + y^2)^{-1/2}(2y)$$

$$= 3xy(x^2 + y^2)^{-1/2}$$

$$\frac{\partial z}{\partial y} = \frac{3}{2}(x^2 + y^2)^{1/2}(2y)$$

$$= 3y(x^2 + y^2)^{1/2}$$

$$\frac{\partial^2 z}{\partial x \partial y} = \frac{3}{2}y(x^2 + y^2)^{-1/2}(2x)$$

$$= 3xy(x^2 + y^2)^{-1/2}$$

$$\frac{\partial^2 z}{\partial y \partial x} = \frac{\partial^2 z}{\partial x \partial y} = 3xy(x^2 + y^2)^{-1/2}$$

EXAMPLE

If $u = Ax^4 + Bx^3y + Cx^2y^2 + Dxy^3 + Ey^4$, find $\frac{\partial^3 u}{\partial x^3}$, $\frac{\partial^3 u}{\partial x \partial y^2}$, $\frac{\partial^3 u}{\partial x^2 \partial y}$, and $\frac{\partial^3 u}{\partial y^3}$.

$$\frac{\partial u}{\partial x} = 4Ax^3 + 3Bx^2y + 2Cxy^2 + Dy^3$$

$$\frac{\partial^2 u}{\partial x^2} = 12Ax^2 + 6Bxy + 2Cy^2$$

$$\frac{\partial^3 u}{\partial x^3} = 24Ax + 6By$$

$$\frac{\partial^3 u}{\partial x^2 \partial y} = 6Bx + 4Cy$$

$$\frac{\partial u}{\partial y} = Bx^3 + 2Cx^2y + 3Dxy^2 + 4Ey^3$$

$$\frac{\partial^2 u}{\partial y^2} = 2Cx^2 + 6Dxy + 12Ey^2$$

$$\frac{\partial^3 u}{\partial y^2 \partial x} = 4Cx + 6Dy$$

$$\frac{\partial^3 u}{\partial y^3} = 6Dx + 24Ey$$

PROBLEMS

1. If $u = xy - \ln xy$, find u_x and u_y .
2. If $z = (x - y)/(x + y)$, find z_x and z_y .

3. If $z = ye^{x/y}$, find z_x and z_y .
4. If $z = \ln \left(\frac{x^2 - y^2}{x^2 + y^2} \right)$, find z_x and z_y .
5. If $u = \sin xy$, find u_x and u_y .
6. If $z = \sqrt{xy}$, find z_x and z_y .
7. If $z = e^{au + bv^2 + cw^3}$, find z_u , z_v , and z_w .
8. If $u = x^2y + y^2z + z^2x$, show that $u_x + u_y + u_z = (x + y + z)^2$.
9. If $z = x^2 \sin(y/x) + y^2 \cos(y/x)$, find $xz_x + yz_y$.
10. If $a = \cos xu \sin yu$, find z_x , z_y , and z_u .
11. If $z = xy + y \ln xy$, show that $x \frac{\partial^2 z}{\partial x^2} + y \frac{\partial^2 z}{\partial x \partial y} = y^2 \frac{\partial^2 z}{\partial y^2}$.
12. If $u = \sqrt{x - y^2}$, show that $\frac{\partial u}{\partial x} \cdot \frac{\partial^2 u}{\partial x \partial y} = \frac{\partial u}{\partial y} \cdot \frac{\partial^2 u}{\partial x^2}$.
13. If $z = xy + xe^{1/y}$, show that $\frac{\partial^2 z}{\partial x \partial y} = \frac{\partial^2 z}{\partial y \partial x}$.
14. If $v = \frac{x + y}{x - y}$, show that $\frac{\partial^2 v}{\partial x \partial y} = \frac{\partial^2 v}{\partial y \partial x}$.
15. If $z = 2x^2 - 2y^2 - 3x - 4xy^2$, show that $\frac{\partial^2 z}{\partial x \partial y} = \frac{\partial^2 z}{\partial y \partial x}$.
16. If $z = \ln(x^2 + y^2)$, show that $\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} = 0$.
17. If $f(x, y) = x^3 e^{x^2 + y}$, find f_x , f_y , and f_{xy} .
18. If $f(x, y) = Ax + By + Ce^{xy}$, find f_{xx} , f_{yy} , and f_{xy} .
19. If $f(x, y) = x^2 \cos y + y^2 \sin x$, find f_{xx} , f_{yy} , and f_{xy} .
20. If $f(x, y) = x^3 + 3x^2y + 6xy^2 - y^3$, find $f_{xx}(2, 3)$, $f_{yy}(2, 3)$, and $f_{xy}(2, 3)$.
21. If $f(x, y) = x^4 - 4x^3y + 8xy^3 - y^4$, find $f_{xx}(0, 1)$, $f_{yy}(0, 1)$, and $f_{xy}(0, 1)$.
22. If $f(x, y) = 2x^2 - 3xy + 4y^2$, find $f_x(1, -1)$ and $f_y(1, -1)$.
23. If $f(x, y) = \frac{2x}{x - y}$, find $f_x(3, 1)$ and $f_y(3, 1)$.
24. If $f(x, y) = e^{-x} \sin(x + 2y)$, find $f_x\left(0, \frac{\pi}{4}\right)$ and $f_y\left(0, \frac{\pi}{4}\right)$.
25. If $u = \frac{x^2y^2}{x + y}$, show that $xu_x + yu_y = 3u$.
26. If $u = \ln(x + y + z)$, show that $\ln u_x + \ln u_y + \ln u_z = -3u$.
27. If $u = \frac{Ax^n + By^n}{Cx^2 + Dy^2}$, show that $xu_x + yu_y = (n - 2)u$.
28. If $u = f(x, y)$, $v = g(x, y)$, $u_x = v_y$, and $u_y = -v_x$, show that if $x = r \cos \theta$ and $y = r \sin \theta$, then $u_r = \frac{1}{r} v_\theta$ and $v_r = \frac{1}{r} u_\theta$.
29. If $u = f(x, y)$, $x = r \cos \theta$, and $y = r \sin \theta$, find $u_r^2 + \frac{1}{r^2} u_\theta^2$ in terms of x and y .

30. If $z = f(x, y)$, $x = g(u, v)$, and $y = h(u, v)$, find z_{uu} .
31. If $u = \ln \sqrt{x^2 + y^2}$, show that $u_{xx} = -u_{yy}$.
32. If $u = \frac{1}{r}$ and $r = \sqrt{x^2 + y^2 + z^2}$, show that $u_{xx} + u_{yy} + u_{zz} = \frac{1}{r^4}$.

Answers to Odd-Numbered Problems

1. $u_x = y - \frac{1}{x}$; $u_y = x - \frac{1}{y}$
3. $z_x = e^{x/y}$; $z_y = e^{x/y} \left(1 - \frac{x}{y}\right)$
5. $u_x = y \cos xy$; $u_y = x \cos xy$
7. $z_u = ae^{au + bv^2 + cw^3}$; $z_v = 2bve^{au + bv^2 + cw^3}$; $z_w = 3cw^2e^{au + bv^2 + cw^3}$
9. $2z$
17. $f_x = x^2(3 + 2x^2)e^{x^2+y}$; $f_y = x^3e^{x^2+y}$; $f_{xy} = x^2(3 + 2x^2)e^{x^2+y}$
19. $f_{xx} = 2 \cos y - y^2 \sin x$; $f_{yy} = -x^2 \cos y + 2 \sin x$;
 $f_{xy} = -2x \sin y + 2y \cos x$
21. $f_{xx}(0, 1) = 0$; $f_{yy}(0, 1) = -12$; $f_{xy}(0, 1) = 24$
23. $f_x(3, 1) = -\frac{1}{2}$; $f_y(3, 1) = \frac{3}{2}$
29. $u_x^2 + u_y^2$

THE TOTAL DIFFERENTIAL

The total differential of a function

$$w = f(x, y, z)$$

is defined by

$$dw = \frac{\partial w}{\partial x} dx + \frac{\partial w}{\partial y} dy + \frac{\partial w}{\partial z} dz$$

The separate terms $\frac{\partial w}{\partial x}$, $\frac{\partial w}{\partial y}$ and $\frac{\partial w}{\partial z}$ are sometimes called partial differentials of w with respect to x , y , and z , respectively. The sum of the *partial differentials* of a function is its *total differential*.

In general, the total differential of a function

$$w = f(x_1, x_2, \dots, x_n)$$

is the sum of all its partial differentials

$$dw = \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i$$

If the x 's are differentiable functions of another variable, say t , then

$$dx_i = \frac{dx_i}{dt} dt$$

and if the x 's are differentiable functions of two variables, say r and s , then

$$dx_i = \frac{\partial x_i}{\partial r} dr + \frac{\partial x_i}{\partial s} ds$$

EXAMPLE

If $z = 2x^3 - 4xy^2 + 3y^3$,

$$\begin{aligned} dz &= 6x^2 dx - 4y^2 dx - 8xy dy + 9y^2 dy \\ &= (6x^2 - 4y^2) dx + (9y^2 - 8xy) dy \end{aligned}$$

EXAMPLE

If $w = x^2 + y^2 + z^2$, where $x = r \cos t$, $y = r \sin t$, and $z = r$, then

$$\begin{aligned} dw &= \frac{\partial w}{\partial x} dx + \frac{\partial w}{\partial y} dy + \frac{\partial w}{\partial z} dz \\ &= 2x dx + 2y dy + 2z dz \end{aligned}$$

But

$$dx = \cos t dr - r \sin t dt$$

$$dy = \sin t dr + r \cos t dt$$

$$dz = dr$$

and thus

$$dw = 2x(\cos t dr - r \sin t dt) + 2y(\sin t dr + r \cos t dt) + 2z dr$$

Substituting for x , y , and z ,

$$\begin{aligned} dw &= 2(r \cos^2 t dr - r^2 \sin t \cos t dt + r \sin^2 t dr + r^2 \sin t \cos t dt + r dr) \\ &= 2(r + r) dr \\ &= 4r dr \end{aligned}$$

EXAMPLE

If $u = (x + y)\sqrt{x - y}$, compute du when $x = 6$, $y = 2$, $dx = \frac{1}{2}$, and $dy = -1$.

$$\begin{aligned} du &= \frac{\partial u}{\partial x} dx + \frac{\partial u}{\partial y} dy \\ &= [(x - y)^{1/2} + (x + y)(\tfrac{1}{2})(x - y)^{-1/2}] dx \\ &\quad + [(x - y)^{1/2} + (x + y)(-\tfrac{1}{2})(x - y)^{-1/2}] dy \\ &= (x - y)^{-1/2}[x - y + \tfrac{1}{2}(x + y)] dx \\ &\quad + (x - y)^{-1/2}[x - y - \tfrac{1}{2}(x + y)] dy \\ &= \tfrac{1}{2}[(x - y)^{-1/2}(3x - y) dx + (x - y)^{-1/2}(x - 3y) dy] \\ &= \tfrac{1}{2}[(\tfrac{1}{2})(16)(\tfrac{1}{2}) + (\tfrac{1}{2})(0)(-1)] \\ &= 2 \end{aligned}$$

THE TOTAL DERIVATIVE

If $w = f(x, y, z)$ has continuous partial derivatives $\frac{\partial w}{\partial x}$, $\frac{\partial w}{\partial y}$, and $\frac{\partial w}{\partial z}$ in some region and x , y , and z are functions of another variable t , then

$$\frac{dw}{dt} = \frac{\partial w}{\partial x} \frac{dx}{dt} + \frac{\partial w}{\partial y} \frac{dy}{dt} + \frac{\partial w}{\partial z} \frac{dz}{dt}$$

and $\frac{dw}{dt}$ is said to be the *total derivative* of w with respect to t . Thus $\frac{dw}{dt}$ represents the rate of change in w as t changes. Heuristically, the total derivative can be obtained by dividing the total differential by Δt and letting Δt approach zero as a limit. Note that w is actually a function only of t .

EXAMPLE

If $w = x^2 + y^2 + z^2$, where $x = e^t \cos t$, $y = e^t \sin t$, and $z = e^t$, find $\frac{dw}{dt}$.

$$\frac{dw}{dt} = \frac{\partial w}{\partial x} \frac{dx}{dt} + \frac{\partial w}{\partial y} \frac{dy}{dt} + \frac{\partial w}{\partial z} \frac{dz}{dt}$$

$$\frac{dw}{dt} = 2x(e^t \cos t - e^t \sin t) + 2y(e^t \sin t + e^t \cos t) + 2ze^t$$

Substituting for x , y , and z ,

$$\begin{aligned} \frac{dw}{dt} &= 2[e^{2t} \cos^2 t - e^{2t} \sin t \cos t + e^{2t} \sin^2 t + e^{2t} \sin t \cos t + e^{2t}] \\ &= 2[e^{2t}(\cos^2 t + \sin^2 t)] + 2e^{2t} \\ &= 4e^{2t} \end{aligned}$$

Alternatively,

$$\begin{aligned} w &= e^{2t} \cos^2 t + e^{2t} \sin^2 t + e^{2t} \\ \frac{dw}{dt} &= 2e^{2t} \cos^2 t - 2e^{2t} \sin t \cos t + 2e^{2t} \sin^2 t + 2e^{2t} \sin t \cos t + 2e^{2t} \\ &= 2e^{2t}(\cos^2 t + \sin^2 t) + 2e^{2t} \\ &= 4e^{2t} \end{aligned}$$

EXAMPLE

If $u = \frac{x}{y} + \frac{y}{z}$, where $x = 2t$, $y = \frac{1}{t}$, and $z = t^2$,

$$\begin{aligned} \frac{du}{dt} &= \frac{\partial u}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial u}{\partial y} \cdot \frac{dy}{dt} + \frac{\partial u}{\partial z} \cdot \frac{dz}{dt} \\ &= \left(\frac{1}{y}\right)(2) + \left(-\frac{x}{y^2} + \frac{1}{z}\right)\left(-\frac{1}{t^2}\right) + \left(-\frac{y}{z^2}\right)(2t) \\ &= \frac{2}{y} + \frac{x}{y^2 t^2} - \frac{1}{z t^2} - \frac{2yt}{z^2} \end{aligned}$$

Substituting for x , y , and z ,

$$\begin{aligned}\frac{du}{dt} &= 2t + 2t - \frac{1}{t^4} - \frac{2}{t^4} \\ &= 4t - \frac{3}{t^4}\end{aligned}$$

Similarly, if $w = f(x, y, z)$ and x , y , and z are differentiable functions of r and s , then $\frac{\partial w}{\partial r}$ and $\frac{\partial w}{\partial s}$ can be obtained as follows:

$$\frac{\partial w}{\partial r} = \frac{\partial w}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial w}{\partial y} \frac{\partial y}{\partial r} + \frac{\partial w}{\partial z} \frac{\partial z}{\partial r}$$

$$\frac{\partial w}{\partial s} = \frac{\partial w}{\partial x} \frac{\partial x}{\partial s} + \frac{\partial w}{\partial y} \frac{\partial y}{\partial s} + \frac{\partial w}{\partial z} \frac{\partial z}{\partial s}$$

In general, if w is a differentiable function of x_1, x_2, \dots, x_n and the x 's are differentiable functions of a second set of variables u_1, u_2, \dots, u_m , then the partial derivative of w with respect to a variable in the second set, say u_j , is given by

$$\frac{\partial w}{\partial u_j} = \frac{\partial w}{\partial x_1} \frac{\partial x_1}{\partial u_j} + \frac{\partial w}{\partial x_2} \frac{\partial x_2}{\partial u_j} + \dots + \frac{\partial w}{\partial x_n} \frac{\partial x_n}{\partial u_j}$$

EXAMPLE

If $u = x^2 - xy + y^2$, where $x = rs$, $y = r^2 + s^2$,

$$\begin{aligned}\frac{\partial u}{\partial r} &= \frac{\partial u}{\partial x} \cdot \frac{\partial x}{\partial r} + \frac{\partial u}{\partial y} \cdot \frac{\partial y}{\partial r} \\ &= (2x - y)s + 2(2y - x)r\end{aligned}$$

Substituting for x and y ,

$$\begin{aligned}\frac{\partial u}{\partial r} &= (2rs - r^2 - s^2)s + 2(2r^2 + 2s^2 - rs)r \\ &= 2rs^2 - r^2s - s^3 + 4r^3 + 4s^2r - 2r^2s \\ &= 4r^3 - 3r^2s + 6rs^2 - s^3\end{aligned}$$

$$\begin{aligned}\frac{\partial u}{\partial s} &= \frac{\partial u}{\partial x} \cdot \frac{\partial x}{\partial s} + \frac{\partial u}{\partial y} \cdot \frac{\partial y}{\partial s} \\ &= (2x - y)r + 2(2y - x)s\end{aligned}$$

Substituting for x and y ,

$$\begin{aligned}\frac{\partial u}{\partial s} &= (2rs - r^2 - s^2)r + 2(2r^2 + 2s^2 - rs)s \\ &= 2r^2s - r^3 - rs^2 + 4r^2s + 4s^3 - 2rs^2 \\ &= 4s^3 - 3rs^2 + 6r^2s - r^3\end{aligned}$$

EXAMPLE

If $w = \ln(x^2 + y^2 + 2z)$, where $x = r + s$, $y = r - s$, and $z = 2rs$,

$$\begin{aligned}\frac{\partial w}{\partial r} &= \frac{\partial w}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial w}{\partial y} \frac{\partial y}{\partial r} + \frac{\partial w}{\partial z} \frac{\partial z}{\partial r} \\ &= \left(\frac{2x}{x^2 + y^2 + 2z} \right) (1) + \left(\frac{2y}{x^2 + y^2 + 2z} \right) (1) + \left(\frac{2}{x^2 + y^2 + 2z} \right) (2s) \\ &= \frac{2(x + y + 2s)}{x^2 + y^2 + 2z} \\ &= \frac{2(r + s + r - s + 2s)}{r^2 + 2rs + s^2 + r^2 - 2rs + s^2 + 4rs} \\ &= \frac{4(r + s)}{2(r + s)^2} \\ &= \frac{2}{r + s}\end{aligned}$$

and

$$\begin{aligned}\frac{\partial w}{\partial s} &= \frac{\partial w}{\partial x} \frac{\partial x}{\partial s} + \frac{\partial w}{\partial y} \frac{\partial y}{\partial s} + \frac{\partial w}{\partial z} \frac{\partial z}{\partial s} \\ &= \left(\frac{2x}{x^2 + y^2 + 2z} \right) (1) + \left(\frac{2y}{x^2 + y^2 + 2z} \right) (-1) + \left(\frac{2}{x^2 + y^2 + 2z} \right) (2r) \\ &= \frac{2(x - y + 2r)}{x^2 + y^2 + 2z} \\ &= \frac{2(r + s - r + s + 2r)}{2(r + s)^2} \\ &= \frac{4(r + s)}{2(r + s)^2} \\ &= \frac{2}{r + s}\end{aligned}$$

The composite function rule, which is valid also for partial derivatives, can be used to obtain the total derivatives of functions whose independent variables are related in various ways.

For example, if $u = f(x, y, z)$, where $x = t$, $y = y(t)$, and $z = z(t)$, then

$$\begin{aligned}\frac{du}{dt} &= \frac{\partial u}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial u}{\partial y} \frac{\partial y}{\partial t} + \frac{\partial u}{\partial z} \frac{\partial z}{\partial t} \\ &= \frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} \frac{\partial y}{\partial t} + \frac{\partial u}{\partial z} \frac{\partial z}{\partial t}\end{aligned}$$

since $\frac{\partial x}{\partial t} = 1$.

If $u = f(x, y, z)$, where $x = x(t)$, $y = y(x)$, and $z = z(t)$, then

$$\frac{du}{dt} = \frac{\partial u}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial u}{\partial y} \frac{\partial y}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial u}{\partial z} \frac{\partial z}{\partial t}.$$

using the composite function rule.

If $u = f(x, y, z)$, where $x = x(r, s)$, $y = y(r)$, and $z = z(y)$, then

$$\frac{\partial u}{\partial r} = \frac{\partial u}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial u}{\partial y} \frac{\partial y}{\partial r} + \frac{\partial u}{\partial z} \frac{\partial z}{\partial y} \frac{\partial y}{\partial r}$$

using the composite function rule, and

$$\frac{\partial u}{\partial s} = \frac{\partial u}{\partial x} \frac{\partial x}{\partial s}$$

since $\frac{\partial y}{\partial s} = 0$.

If $u = f(x, y, z)$, where $x = x(t)$, $y = y(x)$, and $z = z(y)$, then

$$\frac{du}{dt} = \frac{\partial u}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial u}{\partial y} \frac{\partial y}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial u}{\partial z} \frac{\partial z}{\partial y} \frac{\partial y}{\partial x} \frac{\partial x}{\partial t}$$

using the composite function rule.

DIFFERENTIATION OF IMPLICIT FUNCTIONS

The process of implicit differentiation discussed in previous sections can be formalized using partial derivatives and can then be generalized to provide formulas for obtaining partial derivatives of implicit functions of several variables.

The equation $f(x, y) = 0$ defines x and y as implicit functions of each other. For purposes of derivation, let

$$u = f(x, y)$$

Then

$$\frac{\partial u}{\partial x} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} \frac{dy}{dx}$$

But if $f(x, y) = 0$, then $u = 0$, $\frac{\partial u}{\partial x} = 0$ and

$$\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} \frac{dy}{dx} = 0$$

Thus

$$\frac{dy}{dx} = - \frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}}$$

and $\frac{dx}{dy}$ can be obtained as the reciprocal of $\frac{dy}{dx}$. Note that the equation

$\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} \cdot \frac{dy}{dx} = 0$ represents the process of implicit differentiation discussed previously.

EXAMPLE

If $e^x \sin y + e^y \cos x = 1$, find $\frac{dy}{dx}$ and $\frac{dx}{dy}$.

$$f(x, y) = e^x \sin y + e^y \cos x - 1 = 0$$

$$\begin{aligned} \frac{dy}{dx} &= -\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}} \\ &= -\frac{e^x \sin y - e^y \sin x}{e^x \cos y + e^y \cos x} \end{aligned}$$

The usual procedure of implicit differentiation can also be used to obtain $\frac{dy}{dx}$, as follows:

$$e^x \sin y + e^x \cos y \frac{dy}{dx} + e^y \cos x \frac{dy}{dx} - e^y \sin x = 0$$

$$\frac{dy}{dx} = \frac{e^y \sin x - e^x \sin y}{e^x \cos y + e^y \cos x}$$

$\frac{dx}{dy}$ can be obtained by either of the methods above or, more easily, as the reciprocal of $\frac{dy}{dx}$.

$$\begin{aligned} \frac{dx}{dy} &= \frac{1}{\frac{dy}{dx}} \\ &= \frac{e^x \cos y + e^y \cos x}{e^y \sin x - e^x \sin y} \end{aligned}$$

EXAMPLE

If $f(x, y) = x^3 + y^3 - 3axy = 0$, find $\frac{dy}{dx}$ and $\frac{dx}{dy}$.

$$\begin{aligned} \frac{dy}{dx} &= -\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}} \\ &= -\frac{3x^2 - 3ay}{3y^2 - 3ax} \\ &= \frac{ay - x^2}{y^2 - ax} \end{aligned}$$

$$\begin{aligned} \frac{dx}{dy} &= \frac{1}{\frac{dy}{dx}} \\ &= \frac{y^2 - ax}{ay - x^2} \end{aligned}$$

Note that $\frac{dy}{dx}$ and $\frac{dx}{dy}$ could also be obtained using the usual procedure of implicit differentiation.

Similar formulas are appropriate for obtaining the partial derivatives of implicit functions of any finite number of variables; these formulas can be obtained using the same method of derivation shown above.

If z is defined as an implicit function of x and y by the equation $F(x, y, z) = 0$, then

$$\frac{\partial z}{\partial x} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial z}} \quad \text{for } \frac{\partial F}{\partial z} \neq 0$$

$$\frac{\partial z}{\partial y} = -\frac{\frac{\partial F}{\partial y}}{\frac{\partial F}{\partial z}} \quad \text{for } \frac{\partial F}{\partial z} \neq 0$$

The choice of z as the dependent variable is arbitrary and other partial derivatives can also be obtained; for example,

$$\frac{\partial y}{\partial x} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}} \quad \text{for } \frac{\partial F}{\partial y} \neq 0$$

similarly, $\frac{\partial x}{\partial z}$, $\frac{\partial y}{\partial z}$, and $\frac{\partial x}{\partial y}$ can be obtained directly or as reciprocals.

More generally, if z is defined as an implicit function of x_1, x_2, \dots, x_n by the equation $F(x_1, x_2, \dots, x_n, z) = 0$, then

$$\frac{\partial z}{\partial x_i} = -\frac{\frac{\partial F}{\partial x_i}}{\frac{\partial F}{\partial z}} \quad \text{for } i = 1, 2, \dots, n \quad \text{and} \quad \frac{\partial F}{\partial z} \neq 0$$

Again, choice of z as the dependent variable is arbitrary and $\frac{\partial x_i}{\partial x_j}$ can also be obtained as follows

$$\frac{\partial x_i}{\partial x_j} = -\frac{\frac{\partial F}{\partial x_j}}{\frac{\partial F}{\partial x_i}} \quad \text{for } i = 1, 2, \dots, n, i \neq j \quad \text{and} \quad \frac{\partial F}{\partial x_i} \neq 0$$

EXAMPLE

Find $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ for the function $ze^x + e^y - ye^z = 0$.

$$\frac{\partial F}{\partial x} = ze^x = -e^y + ye^z$$

$$\frac{\partial F}{\partial y} = e^y - e^z$$

$$\frac{\partial F}{\partial z} = e^x - ye^z$$

$$\frac{\partial z}{\partial x} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial z}} = \frac{e^y - ye^z}{e^x - ye^z}$$

$$\frac{\partial z}{\partial y} = -\frac{\frac{\partial F}{\partial y}}{\frac{\partial F}{\partial z}} = \frac{e^z - e^y}{e^x - ye^z}$$

EXAMPLE

Find $\frac{\partial z}{\partial x}$ and $\frac{\partial y}{\partial x}$ for the function $z \sin x + z^2 + x \sin y + xyz = 0$.

$$\frac{\partial z}{\partial x} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial z}} = -\frac{z \cos x + \sin y + yz}{\sin x + 2z + xy}$$

$$\frac{\partial y}{\partial x} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}} = -\frac{z \cos x + \sin y + yz}{x \cos y + xz}$$

PROBLEMS

1. If $z = x^3 + x^2y - y^3$, find dz .
2. If $u = \ln(x^2 + y^2 + z^2)^{1/2}$, find du .
3. If $u = e^{xyz}$, find du .
4. If $u = e^z \sin[(x - y)z]$, find du .
5. If $z = 2x^3 - 4xy^2 + 3y^3$, find dz .
6. If $u = xy^2z^3$, find du .
7. If $x^2 + y^2 + z^2 = a^2$, find dz .
8. If $u = x + 4x^{1/2}y^{1/2} - 3y$, $x = t^3$, and $y = \frac{1}{t}$, find $\frac{du}{dt}$.
9. If $x^3 + y^3 - 3bxy = 0$, find $\frac{dy}{dx}$.
10. If $x^2 + 2xy + 2y = 15$, find $\frac{dy}{dx}$ if $x = 2$, $y = 3$.
11. If $x^3 - y^3 - 4xy = -\frac{1}{2}$, find $\frac{dy}{dx}$ if $x = 2$, $y = -2$.
12. If $Ax + By + Ce^{xy} = D$, find $\frac{dy}{dx}$ if $x = y = 0$.
13. If $Ax^2 + By^2 + Cz^2 = D$, find z_x and z_y .
14. If $xy + yz + zx = 9xyz$, find z_x and z_y .

15. If $xz = \cos yz + a$, find z_x and z_y .
16. If $e^x + e^y + e^z = axyz$, find y_x .
17. If $F(x, y, z) = 0$, show that $x_y y_z z_x = -1$.
18. If $u = x^3 - 3xy + y^3$, $x = r^2 + s$, and $y = rs^2$, find $\frac{\partial u}{\partial r}$.
19. If $u = xy + yz$, $x = \frac{e^t}{t}$, $y = \frac{e^{-t}}{t}$, and $z = t^2$, find $\frac{du}{dt}$.
20. If $z = x \ln y + y \ln x$, $x = e^{u+v}$, and $y = e^{u-v}$, find $\frac{\partial z}{\partial u}$ and $\frac{\partial z}{\partial v}$.
21. If $z = \ln(x^2 + y^2) + \sqrt{x^2 + y^2}$, $x = e^u \cos v$, and $y = e^u \sin v$, find $\frac{\partial z}{\partial u}$.
22. If $u = xy + yz + zx$, $x = r^s$, $y = sr$, and $z = r + s$, find u_r .
23. If $x \ln yz - y \ln xz = 0$, find z_x and z_y .
24. If $e^{xyz} = e^x + e^y + e^z$, find z_x and z_y .
25. If $e^x + e^y + e^z = e^{x+y+z}$, find z_x and z_y .
26. If $u \ln \frac{v}{w} - w \ln uv = 0$, find u_v and u_w .

Answers to Odd-Numbered Problems

1. $dz = (3x^2 + 2xy) dx + (x^2 - 3y^2) dy$
3. $du = e^{xyz}(yz dx + xz dy + xy dz)$
5. $dz = (6x^2 - 4y^2) dx + (9y^2 - 8xy) dy$
7. $dz = -(x dx + y dy)/z$
9. $\frac{dy}{dx} = \frac{by - x^2}{y^2 - bx}$
11. $\frac{dy}{dx} = 1$
13. $z_x = -\frac{Ax}{Cz}$; $z_y = -\frac{By}{Cz}$
15. $z_x = -z/(x + y \sin yz)$; $z_y = -(z \sin yz)/(x + y \sin yz)$
19. $\frac{du}{dt} = e^{-t}(1 - t) - \frac{2}{t^3}$
21. $\frac{\partial z}{\partial u} = 2 + e^u$
23. $z_x = \frac{yz(1 - \ln xz)}{x(x - y)}$
 $z_y = \frac{xz(1 - \ln yz)}{y(y - x)}$
25. $z_x = -\frac{e^y + e^z}{e^x + e^y}$
 $z_y = -\frac{e^x + e^z}{e^x + e^y}$

3.3 Applications of Partial Derivatives in Business and Economics

In this section several applications of partial derivatives in business and economics are discussed. These applications include marginal cost, demand surfaces, production functions, Euler's theorem, constant product curves, returns to scale, and utility functions.

MARGINAL COST

If the joint-cost function for producing the quantities x and y of two commodities is given by

$$C = Q(x, y)$$

then the partial derivatives of C are the *marginal cost functions*:

$$\frac{\partial C}{\partial x} \text{ is the marginal cost with respect to } x$$

$$\frac{\partial C}{\partial y} \text{ is the marginal cost with respect to } y$$

In most economic situations marginal costs are positive.

EXAMPLE

If the joint-cost function for producing quantities x and y of two commodities is

$$C = x \ln(5 + y)$$

then

$$\frac{\partial C}{\partial x} = \ln(5 + y) \text{ is the marginal cost with respect to } x$$

and

$$\frac{\partial C}{\partial y} = \frac{x}{5 + y} \text{ is the marginal cost with respect to } y$$

EXAMPLE

If the joint-cost function for producing quantities x and y of two commodities is

$$C = 15 + 2x^2 + xy + 5y^2$$

then

$$\frac{\partial C}{\partial x} = 4x + y \text{ is the marginal cost with respect to } x$$

and

$$\frac{\partial C}{\partial y} = x + 10y \text{ is the marginal cost with respect to } y$$

If $x = 3$ and $y = 6$, $\frac{\partial C}{\partial x} = 18$ and $\frac{\partial C}{\partial y} = 63$. Thus, if y is kept constant at 6, producing an additional unit of x adds \$18 to the total cost; if x is kept constant at 3, producing an additional unit of y adds \$63 to the total cost.

DEMAND SURFACES

If there are two related commodities for which the quantities demanded are x and y and the respective prices are p and q , then the demand functions can be represented by

$$x = f(p, q) \quad \text{and} \quad y = g(p, q)$$

assuming that the quantities demanded, x and y , depend only on the prices, p and q , of the two commodities. If a demand function of two independent variables is continuous, it can be represented by a surface, referred to as a *demand surface*.

NOTE: The notation p_x and p_y is frequently used for the prices of the two commodities. It is not used in this section to avoid possible confusion with the subscript notation for partial derivatives.

In the usual economic situations, the demand functions $x = f(p, q)$ and $y = g(p, q)$ have the following properties:

1. All the variables, x , y , p , and q , are zero or positive.
2. If q is constant, x is a monotonically decreasing function of p ; similarly, if p is constant, y is a monotonically decreasing function of q .
3. The functions $f(x, y)$ and $g(x, y)$ and the region for which they are defined are such that it is possible to obtain their inverse functions $p = F(x, y)$ and $q = G(x, y)$.

For a constant price p , as q increases y decreases, but x may either increase or decrease; if x increases, the two commodities are said to be *competitive*, since a decrease in the demand for one corresponds to an increase in the demand for the other. For a constant price p , as q decreases, y increases; if x also increases, the commodities are said to be *complementary*, since an increase in the demand for one corresponds to an increase in the demand for the other. Corresponding relationships hold for changes in p for a constant q .

NOTE: Usually, it is assumed that related commodities which are either competitive or complementary at one set of prices have the same relationship at other prices; however, in some cases the relationship between commodities may differ for different prices and thus must be defined for a particular set of prices.

Marginal Demand

If the demand functions for two related commodities are

$$x = f(p, q) \quad \text{and} \quad y = g(p, q)$$

then the partial derivatives of x and y are the *marginal demand functions*:

$$\frac{\partial x}{\partial p} \quad \text{is the marginal demand of } x \text{ with respect to } p$$

$\frac{\partial x}{\partial q}$ is the marginal demand of x with respect to q

$\frac{\partial y}{\partial p}$ is the marginal demand of y with respect to p

$\frac{\partial y}{\partial q}$ is the marginal demand of y with respect to q

For the usual demand functions, x increases if its corresponding price p decreases and y increases if its corresponding price q decreases, so $\frac{\partial x}{\partial p}$ and $\frac{\partial y}{\partial q}$ are negative for all economically meaningful values of p and q .

If $\frac{\partial x}{\partial q}$ and $\frac{\partial y}{\partial p}$ are both negative for given (p, q) , the commodities are complementary, since then a decrease in either price corresponds to increases in both demands. If $\frac{\partial x}{\partial q}$ and $\frac{\partial y}{\partial p}$ are both positive for given (p, q) , the goods are competitive, since then a decrease in either price corresponds to an increase in one demand and a decrease in the other. If $\frac{\partial x}{\partial q}$ and $\frac{\partial y}{\partial p}$ have opposite signs, the commodities are neither complementary nor competitive; in this case, a decrease in the price of one of the commodities corresponds to increases in both demands while a decrease in the price of the other commodity corresponds to an increase in one demand and a decrease in the other. Such a situation is unusual but might occur, for example, if two different grades of a material could be used and the better grade material could be obtained only by processing the lower grade material.

NOTE: Complementary and competitive commodities are defined in terms of the signs of their cross elasticities in Chapter 2. Those definitions and the definitions above are equivalent, since the signs of the cross elasticities are the same as the signs of the appropriate partial derivatives representing marginal demand.

EXAMPLE

If the demand surfaces are linear functions of p and q ,

$$x = a_1 + b_1 p + c_1 q \quad y = a_2 + b_2 p + c_2 q$$

then the marginal demand functions are

$$\frac{\partial x}{\partial p} = b_1 \quad \frac{\partial y}{\partial p} = b_2$$

$$\frac{\partial x}{\partial q} = c_1 \quad \frac{\partial y}{\partial q} = c_2$$

Thus for the usual economic situations, $b_1 < 0$ and $c_2 < 0$. If c_1 and b_2 are both positive, the commodities are competitive; if c_1 and b_2 are both negative, the commodities are complementary.

EXAMPLE

If the demand functions for two related commodities are given by

$$x = \frac{a}{p^2 q} \quad y = \frac{a}{p q} \quad a > 0$$

then the marginal demand functions are

$$\frac{\partial x}{\partial p} = -\frac{2a}{p^3q} \quad \frac{\partial y}{\partial p} = -\frac{a}{p^2q}$$

$$\frac{\partial x}{\partial q} = -\frac{a}{p^2q^2} \quad \frac{\partial y}{\partial q} = -\frac{a}{pq^2}$$

Since $\frac{\partial x}{\partial q} < 0$ and $\frac{\partial y}{\partial q} < 0$, the commodities are complementary.

EXAMPLE

If the demand functions for two related commodities are given by

$$x = ae^{q-p} \quad y = be^{p-q} \quad a > 0, b > 0$$

then the marginal demand functions are

$$\frac{\partial x}{\partial p} = -ae^{q-p} \quad \frac{\partial y}{\partial p} = be^{p-q}$$

$$\frac{\partial x}{\partial q} = ae^{q-p} \quad \frac{\partial y}{\partial q} = -be^{p-q}$$

Since $\frac{\partial x}{\partial q} > 0$ and $\frac{\partial y}{\partial p} > 0$, the commodities are competitive.

EXAMPLE

If the demand functions for two related commodities are given by

$$x = a_1q^2 - b_1pq \quad y = a_2p^2 - b_2pq \quad a_1 > 0, a_2 > 0, b_1 > 0, b_2 > 0$$

then the marginal demand functions are

$$\frac{\partial x}{\partial p} = -b_1q \quad \frac{\partial y}{\partial p} = 2a_2p - b_2q$$

$$\frac{\partial x}{\partial q} = 2a_1q - b_1p \quad \frac{\partial y}{\partial q} = -b_2p$$

$$\frac{\partial x}{\partial q} > 0 \text{ if } \frac{q}{p} > \frac{b_1}{2a_1} \quad \frac{\partial y}{\partial p} > 0 \text{ if } \frac{q}{p} < \frac{2a_2}{b_2}$$

$$\frac{\partial x}{\partial q} < 0 \text{ if } \frac{q}{p} < \frac{b_1}{2a_1} \quad \frac{\partial y}{\partial p} < 0 \text{ if } \frac{q}{p} > \frac{2a_2}{b_2}$$

$$x > 0 \text{ if } \frac{q}{p} > \frac{b_1}{a_1} \quad y > 0 \text{ if } \frac{q}{p} < \frac{a_2}{b_2}$$

Thus the demand functions are appropriate for values of p and q such that

$$\frac{b_1}{a_1} < \frac{q}{p} < \frac{a_2}{b_2}$$

and the commodities are competitive for values of p and q such that the demand functions are appropriate (since $b_1/a_1 < q/p < a_2/b_2 \Rightarrow b_1/2a_1 < q/p < 2a_2/b_2$).

EXAMPLE

If the demand functions for two related commodities are given by

$$x = ae^{-pq} \quad y = be^{p-q} \quad a > 0, b > 0$$

then the marginal demand functions are

$$\begin{aligned} \frac{\partial x}{\partial p} &= -aqe^{-pq} & \frac{\partial y}{\partial p} &= be^{p-q} \\ \frac{\partial x}{\partial q} &= -ape^{-pq} & \frac{\partial y}{\partial q} &= -be^{p-q} \end{aligned}$$

Since $\frac{\partial x}{\partial q} < 0$ and $\frac{\partial y}{\partial p} > 0$, the commodities are neither competitive nor complementary.

Partial Elasticities of Demand

If the demand functions for two related commodities are

$$x = f(p, q) \quad \text{and} \quad y = g(p, q)$$

then the partial elasticities of demand are given by

$$\left. \frac{Ex}{Ep} \right|_{q=c_1} = \frac{p}{x} \cdot \frac{\partial x}{\partial p} = \frac{\frac{\partial}{\partial x} \ln x}{\frac{\partial}{\partial x} \ln p} \quad \text{the partial elasticity of demand } x \text{ with respect to price } p, \text{ for a constant price } q = c_1$$

$$\left. \frac{Ex}{Eq} \right|_{p=c_2} = \frac{q}{x} \cdot \frac{\partial x}{\partial q} = \frac{\frac{\partial}{\partial x} \ln x}{\frac{\partial}{\partial x} \ln q} \quad \text{the partial elasticity of demand } x \text{ with respect to price } q, \text{ for a constant price } p = c_2$$

$$\left. \frac{Ey}{Ep} \right|_{q=c_3} = \frac{p}{y} \cdot \frac{\partial y}{\partial p} = \frac{\frac{\partial}{\partial y} \ln y}{\frac{\partial}{\partial y} \ln p} \quad \text{the partial elasticity of demand } y \text{ with respect to price } p, \text{ for a constant price } q = c_3$$

$$\left. \frac{Ey}{Eq} \right|_{p=c_4} = \frac{q}{y} \cdot \frac{\partial y}{\partial q} = \frac{\frac{\partial}{\partial y} \ln y}{\frac{\partial}{\partial y} \ln q} \quad \text{the partial elasticity of demand } y \text{ with respect to price } q, \text{ for a constant price } p = c_4$$

Note that Ex/Eq and Ey/Ep are cross partial elasticities of demand; their signs, which are the same as the signs of the corresponding marginal demands, can be used to determine whether commodities are competitive or complementary.

EXAMPLE

If the demand functions for two related commodities are given by

$$x = ae^{q-p} \quad y = be^{p-q}$$

then the partial elasticities of demand are

$$\frac{E_x}{E_p} = \frac{p}{x} \cdot \frac{\partial x}{\partial p} = \frac{p}{ae^{q-p}} (-ae^{q-p}) = -p$$

$$\frac{E_y}{E_p} = \frac{p}{y} \cdot \frac{\partial y}{\partial p} = \frac{p}{be^{p-q}} (be^{p-q}) = p$$

$$\frac{E_x}{E_q} = \frac{q}{x} \cdot \frac{\partial x}{\partial q} = \frac{q}{ae^{q-p}} (ae^{q-p}) = q$$

$$\frac{E_y}{E_q} = \frac{q}{y} \cdot \frac{\partial y}{\partial q} = \frac{q}{be^{p-q}} (-be^{p-q}) = -q$$

and, since $E_x/E_q > 0$ and $E_y/E_p > 0$, x and y are competitive commodities (as above).

EXAMPLE

If the demand functions for two related commodities are given by

$$x = a_1 q^2 - b_1 p q \quad y = a_2 p^2 - b_2 p q \quad a_1 > 0, a_2 > 0, b_1 > 0, b_2 > 0$$

then the partial elasticities of demand are

$$\frac{E_x}{E_p} = \frac{p}{x} \cdot \frac{\partial x}{\partial p} = \frac{p}{a_1 q^2 - b_1 p q} (-b_1 q) = -\frac{b_1 p}{a_1 q - b_1 p}$$

$$\frac{E_x}{E_q} = \frac{q}{x} \cdot \frac{\partial x}{\partial q} = \frac{q}{a_1 q^2 - b_1 p q} (2a_1 q - b_1 p) = \frac{2a_1 q - b_1 p}{a_1 q - b_1 p}$$

$$\frac{E_y}{E_p} = \frac{p}{y} \cdot \frac{\partial y}{\partial p} = \frac{p}{a_2 p^2 - b_2 p q} (2a_2 p - b_2 q) = \frac{2a_2 p - b_2 q}{a_2 p - b_2 q}$$

$$\frac{E_y}{E_q} = \frac{q}{y} \cdot \frac{\partial y}{\partial q} = \frac{q}{a_2 p^2 - b_2 p q} (-b_2 p) = -\frac{b_2 q}{a_2 p - b_2 q}$$

and, since $E_x/E_q > 0$ and $E_y/E_p > 0$, x and y are competitive commodities (as above).

PROBLEMS

1. For the following joint-cost functions, determine the marginal cost with respect to x and the marginal cost with respect to y .

a. $C = x^2 \ln(y + 10)$

c. $C = e^x + e^y + xy + 5$

b. $C = x^3 + 2y^2 - xy + 20$

d. $C = x^2 y^2 - 3xy + y + 8$

2. For each of the following pairs of demand functions, determine the four marginal demands, the nature of the relationship between the two commodities, and the four partial elasticities of demand. Quantities are denoted by x and y and the corresponding prices are denoted by p and q , respectively.

a. $x = 20 - 2p - q$

$y = 9 - p - 2q$

e. $x = \frac{4}{p^2 q}$

$y = \frac{16}{pq^2}$

b. $x = 15 - 2p + q$

$y = 16 + p - q$

f. $x = \frac{4}{pq}$

$y = \frac{16}{pq}$

c. $x = 5 - 2p + q$

$y = 8 - 2p - 3q$

d. $x = \frac{q}{p}$

$y = \frac{p^2}{q}$

Answers to Odd-Numbered Problems

1. a. $\frac{\partial C}{\partial x} = 2x \ln(y + 10)$ c. $\frac{\partial C}{\partial x} = e^x + y$
 $\frac{\partial C}{\partial y} = \frac{x^2}{y + 10}$ $\frac{\partial C}{\partial y} = e^y + x$
 b. $\frac{\partial C}{\partial x} = 3x^2 - y$ d. $\frac{\partial C}{\partial x} = 2xy^2 - 3y$
 $\frac{\partial C}{\partial y} = 4y - x$ $\frac{\partial C}{\partial y} = 2x^2y - 3x + 1$

PRODUCTION FUNCTIONS

The production of most commodities requires the use of at least two factors of production, for example, labor, land, capital, materials, or machines. If the quantity z of a commodity is produced using the amounts x and y , respectively, of two factors of production, then the *production function*

$$z = f(x, y)$$

gives the amount of *output* z when the amounts x and y , respectively, of the *inputs* are used simultaneously. For such a representation to be economically meaningful, it is assumed that the amounts of the inputs can be varied without restriction, at least in the range of interest, and that the production function is continuous.

Marginal Productivity

If the production function is given by $z = f(x, y)$, then the partial derivative $\frac{\partial z}{\partial x}$ of z with respect to x (with y held constant) is the *marginal productivity of x* or the *marginal product of x* ; the partial derivative $\frac{\partial z}{\partial y}$ of z with respect to y (with x held constant) is the *marginal productivity of y* or the *marginal product of y* . Note that the marginal productivity of either input is the rate of increase of the total product as that input is increased, assuming that the amount of the other input remains constant.

Marginal productivity usually is positive for a considerable range, that is, as the amount of one input increases, with the amount of the other input held constant, the output also increases. However, as the input of one factor increases with the input of the other factor held constant, the output usually increases at a decreasing rate until the point is reached at which there is not further increase in output and, in fact, a decrease in output occurs with additional inputs of the factor. This characteristic behavior of production functions is known as the *law of eventually diminishing marginal productivity*.

EXAMPLE

If the production function is given by

$$z = 4x^{3/4}y^{1/4}$$

then the marginal productivity of x is

$$\frac{\partial z}{\partial x} = 3x^{-1/4}y^{1/4}$$

and the marginal productivity of y is

$$\frac{\partial z}{\partial y} = x^{3/4} y^{-3/4}$$

Note that $\frac{\partial z}{\partial x}$ is always positive, but decreases as x increases; similarly, $\frac{\partial z}{\partial y}$ is always positive, but decreases as y increases.

EXAMPLE

If the production function is given by

$$z = 4xy - x^2 - 3y^2$$

then the marginal productivity of x is

$$\frac{\partial z}{\partial x} = 4y - 2x$$

and the marginal productivity of y is

$$\frac{\partial z}{\partial y} = 4x - 6y$$

Note that $\frac{\partial z}{\partial x} > 0$ for $x < 2y$, $\frac{\partial z}{\partial x} = 0$ for $x = 2y$, and $\frac{\partial z}{\partial x} < 0$ for $x > 2y$; similarly, $\frac{\partial z}{\partial y} > 0$ for $y < \frac{2}{3}x$, $\frac{\partial z}{\partial y} = 0$ for $y = \frac{2}{3}x$, and $\frac{\partial z}{\partial y} < 0$ for $y > \frac{2}{3}x$. Thus the marginal productivities at first increase and then decrease as input increases.

EXAMPLE

If a production function is given by

$$z^2 + 4x^2 + 5y^2 - 12xy = 0$$

where z is the amount of output and x and y are the amounts of the inputs, find the marginal productivities.

$$\frac{\partial F}{\partial x} = 8x - 12y$$

$$\frac{\partial F}{\partial y} = 10y - 12x$$

$$\frac{\partial F}{\partial z} = 2z$$

The marginal productivity of x is

$$\frac{\partial z}{\partial x} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial z}} = \frac{6y - 4x}{z}$$

and the marginal productivity of y is

$$\frac{\partial z}{\partial y} = -\frac{\frac{\partial F}{\partial y}}{\frac{\partial F}{\partial z}} = \frac{6x - 5y}{z}$$

Since z is positive for any economically meaningful region of the production function, $\frac{\partial z}{\partial x}$ is positive for $x < \frac{3}{2}y$ and negative for $x > \frac{3}{2}y$ and $\frac{\partial z}{\partial y}$ is positive for $y < \frac{6}{5}x$ and negative for $y > \frac{6}{5}x$.

EULER'S THEOREM

If the function $z = f(x, y)$ has the property that, for any constant λ ,

$$f(\lambda x, \lambda y) = \lambda^n f(x, y)$$

then z is said to be *homogeneous of degree n* . If $n > 0$, the function is said to be *positively homogeneous*; if $n = 1$, the function is said to be *linear homogeneous*.

If $z = f(x, y)$ is positively homogeneous of degree n and the first-order partial derivatives exist, then it can be shown that

$$x \frac{\partial z}{\partial x} + y \frac{\partial z}{\partial y} = n f(x, y)$$

This relationship is known as *Euler's theorem*.

EXAMPLE

The function $z = f(x, y) = 3x^4 + 2x^2y^2 + 7y^4$ is homogeneous of degree 4, since

$$\begin{aligned} f(\lambda x, \lambda y) &= 3\lambda^4 x^4 + 2\lambda^4 x^2 y^2 + 7\lambda^4 y^4 \\ &= \lambda^4 f(x, y) \end{aligned}$$

Thus, according to Euler's theorem,

$$x \frac{\partial z}{\partial x} + y \frac{\partial z}{\partial y} = 4f(x, y)$$

which can be verified as follows:

$$\begin{aligned} x \frac{\partial z}{\partial x} + y \frac{\partial z}{\partial y} &= x(12x^3 + 4xy^2) + y(4x^2y + 28y^3) \\ &= 12x^4 + 8x^2y^2 + 28y^4 \\ &= 4f(x, y) \end{aligned}$$

EXAMPLE

The function $z = f(x, y) = \frac{xy}{x^2 + y^2}$ is homogeneous of degree zero, since

$$\begin{aligned} f(\lambda x, \lambda y) &= \frac{\lambda^2 xy}{\lambda^2(x^2 + y^2)} \\ &= \lambda^0 f(x, y) \end{aligned}$$

Thus, according to Euler's theorem,

$$x \frac{\partial z}{\partial x} + y \frac{\partial z}{\partial y} = 0$$

which can be verified as follows:

$$\begin{aligned} x \frac{\partial z}{\partial x} + y \frac{\partial z}{\partial y} &= x \cdot \frac{y(x^2 + y^2) - xy(2x)}{(x^2 + y^2)^2} + y \cdot \frac{x(x^2 + y^2) - xy(2y)}{(x^2 + y^2)^2} \\ &= \frac{2xy(x^2 + y^2) - 2x^3y - 2xy^3}{(x^2 + y^2)^2} \\ &= 0 \end{aligned}$$

Linear Homogeneous Production Functions

In economic theory, production functions are frequently assumed to be linear homogeneous, because such functions have convenient properties. If a production function $z = f(x, y)$ is linear homogeneous, then according to Euler's theorem,

$$\underbrace{x \frac{\partial z}{\partial x}}_{\substack{\text{total due} \\ \text{to} \\ \text{factor } x}} + \underbrace{y \frac{\partial z}{\partial y}}_{\substack{\text{total due} \\ \text{to} \\ \text{factor } y}} = f(x, y)$$

That is, total production or output is equal to the product of the amount of one input and its marginal productivity plus the product of the amount of the other input and its marginal productivity. The total product is thus allocated to the two factors on the basis of their marginal productivities.

Euler's theorem plays an important part in the development of the marginal-productivity theory of distribution. The basic assumptions of this theory are that each input is paid the value of its marginal productivity and that total output is just exhausted. As shown by Euler's theorem, these conditions are satisfied by linear homogeneous production functions; thus in the development of the marginal-productivity theory of distribution, production functions are generally assumed to be linear homogeneous. Note that if the production function is homogeneous of degree n and each input is paid the value of its marginal productivity, then total output exceeds payments for $n > 1$ and is less than payments for $n < 1$.

CONSTANT PRODUCT CURVES

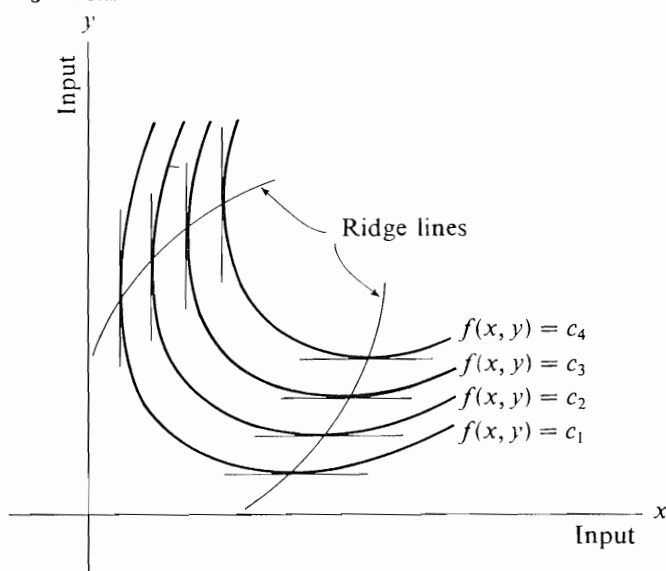
The production function $z = f(x, y)$ is frequently studied by considering the family of curves

$$f(x, y) = \text{constant}$$

in the xy -plane. These curves, no two of which intersect, are called *constant product curves*, *equal product curves*, or *isoquants*. Each curve shows the combinations of factors x and y which result in a specified output. As discussed below, constant product curves are such that a decrease in the input of one factor is compensated for

by an increase in the input of the other factor. The slope $\frac{dy}{dx}$ of the tangent at a point on an isoquant is the rate at which x must be substituted for y to maintain the level of output. The negative of this derivative $-\frac{dy}{dx}$ is defined as the *rate of technical substitution*.

Figure 3.2



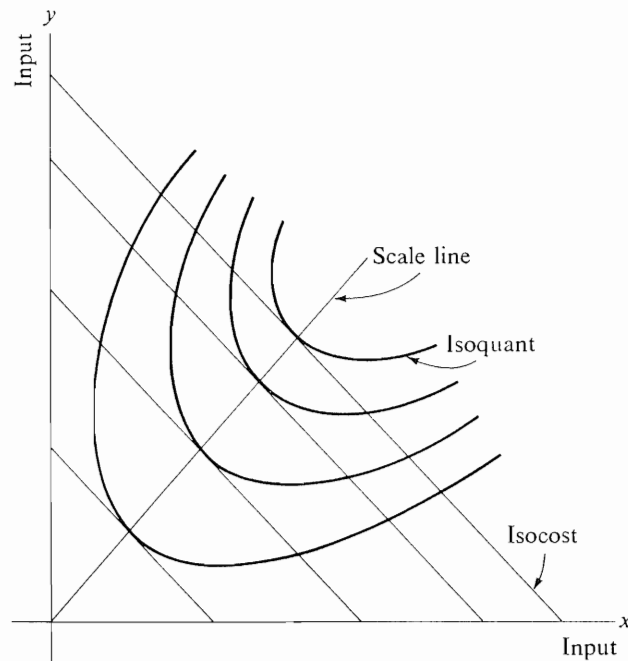
Equal product curves are of various shapes, depending on the nature of the production function, but in general they have negative slopes and are convex to the origin in the area of interest. However, there are some economic situations in which portions of the equal product curves have positive slopes, that is, there are some situations in which marginal productivity is diminishing and additional amounts of a factor of input result in decreased production. For example, if the two factors of production are land and labor, so many men might be employed (for a fixed amount of land) that they would interfere with each other and production would decrease. However, assuming the objective of profit maximization, portions of equal product curves having positive slopes would not be used in practice, since the given output could be obtained at less cost by a factor combination on a portion of the curve having negative slope. Thus, as shown in Figure 3.2, only the portions of the equal product curves lying between their respective vertical and horizontal tangents are relevant; the lines through these points of tangency are known as *ridge lines*. The firm would never operate above the upper ridge line or to the right of the lower ridge line, since this would require more input than necessary to produce a specified output.

The assumption that equal product curves are convex to the origin implies that the marginal significance of one factor with respect to the other factor decreases along an equal product curve—that is, as the amount of factor x that is used increases, the amount of factor y that can be given up in exchange for a further unit of factor x decreases, if the product remains constant; similarly, as the amount of factor y that is used increases, the amount of factor x that can be given up in exchange for a further unit of factor y decreases, if the product remains constant. If the marginal significance of a factor increased as the amount used increased, as would be the case for a concave equal product curve, then only that one factor should be used, since each unit of the factor would be progressively more worth purchasing. Thus, in practice, any portions of equal product curves for which marginal significance of the factors is not decreasing are not relevant.

If factors x and y are purchased in a perfectly competitive market at constant unit prices, then the total cost of production is given by

$$C = p_x x + p_y y + b$$

Figure 3.3



where p_x and p_y are the respective prices of x and y and b is the cost of the fixed inputs. *Isocost curves* give the combinations of input factors that can be purchased for a specified total cost; for the cost function above, the isocost curves are represented by the family of lines

$$p_x x + p_y y + b = \text{constant}$$

The slopes of the isocost lines are equal to the negative of the input price ratio, since

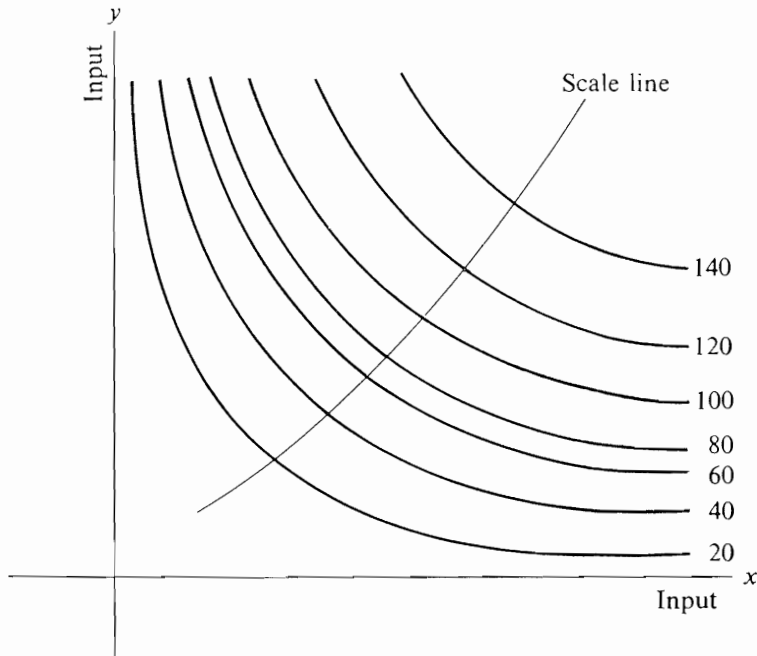
$$y = -\frac{p_x}{p_y}x - \frac{b}{p_y} + \frac{\text{constant}}{p_y}$$

The greatest total output which can be produced for a specified total cost is given by the point of tangency of the relevant isocost curve and an isoquant. Similarly, the lowest possible cost of producing a specified output is given by the point of tangency of the relevant isoquant and an isocost curve. The curve connecting the points of tangency of the isocost curves and isoquants is known as the *expansion path* or *scale line*. This line gives the lowest-cost combinations of factors as output increases; the firm will thus always produce at some point on the scale line (see Figure 3.3). The expansion path, at all points of which the rate of technical substitution equals the fixed input-price ratio, is a straight line if the production function is homogeneous of any degree. A straight-line expansion path does not, however, imply a homogeneous production function.

RETURNS TO SCALE

Returns to scale describes the change in output with a proportionate increase in all inputs. If output increases by the same proportion as the inputs increase, then returns to scale are constant; if output increases by a greater proportion than inputs, returns to scale are increasing; if output increases by a smaller proportion than inputs,

Figure 3.4



returns to scale are decreasing. Very frequently, there are increasing returns to scale for relatively small inputs (economies of scale), returns to scale are then constant for a range of inputs, and finally there are decreasing returns to scale for large inputs (diseconomies of scale).

Returns to scale are easily determined for homogeneous production functions. If a production function is homogeneous of degree n , returns to scale are increasing if $n > 1$, constant if $n = 1$, and decreasing if $n < 1$. In practice, a degree of homogeneity other than 1 is seldom assumed for a production function. If a set of equal product curves (not necessarily homogeneous) represent successive increases in production of a given amount, then returns to scale increase as output increases if the distance along the scale line between successive curves decreases; similarly, returns to scale decrease as output increases if the distance along the scale line between successive curves increases. Thus, in Figure 3.4, returns to scale increase from 20 to 80 units of production and decrease from 80 to 140 units of production.

Note that if the proportions of the amounts of the two factors of production vary, it is necessary to refer to *returns to outlay* rather than returns to scale, since the factors are not increased by the same scale.

UTILITY FUNCTIONS

In the theory of consumer behavior, the consumer is assumed to choose among the alternatives available to him in such a way that the satisfaction derived from consuming commodities is maximized. This implies that the consumer is aware of his alternatives and is capable of evaluating them. All relevant information concerning the satisfaction a consumer derives from various quantities of commodities is contained in his utility function.

The concept of utility is based not on some type of psychological satisfaction on the part of the consumer, but on his choice behavior. The assertion that a consumer derives more satisfaction or utility from A than from B means that he would choose

A rather than B if A and B were presented as alternatives. In addition, the consumer is assumed to make consistent choices so that if A is preferred to B and B is preferred to C , then A is preferred to C .

Modern utility theory assumes only an ordinal utility scale; that is, for all pairs of alternatives A and B , the consumer is assumed to know whether he prefers A to B , prefers B to A , or is indifferent between them; he is not assumed capable of assigning numbers representing amounts of utility to alternatives.

A consumer's ranking (preference order) of commodities is expressed mathematically by his utility function. The utility function associates certain numbers (utility) with various quantities of commodities consumed, but these numbers represent only a ranking or ordering of preferences. Thus if the utility of alternative A is 20 and the utility of alternative B is 5, then A is preferred to B but it is incorrect to say that A is preferred 4 times as strongly as B .

Consider the simple case in which the consumer's purchases are limited to two commodities and the utility function is

$$U = f(q_1, q_2)$$

where q_1 and q_2 are the quantities of the two commodities Q_1 and Q_2 which are consumed. It is assumed that $f(q_1, q_2)$ is continuous and has continuous first-order and second-order partial derivatives.

The utility function is with respect to consumption during a specified period of time, a period long enough for the desire for variety to be satisfied but not so long that tastes change. Utility is derived from the consumption of commodities. By definition, a commodity is something of which the consumer would rather have more than less; otherwise a discommodity is involved. In reality, a commodity consumed in sufficiently large amounts may become a discommodity (for example, too much candy); the following discussion assumes that such a point of satiation has not been reached. Thus the partial derivatives of U with respect to q_1 and q_2 , denoted f_1 and f_2 , are positive.

Since the utility function is continuous, a given level of utility can be derived from an infinite number of combinations of q_1 and q_2 . The locus of all combinations of quantities of commodities from which the consumer derives the same utility is an *indifference curve*. A collection of indifference curves corresponding to different levels of utility is an *indifference map*. (See Figure 3.5.) One indifference curve passes through every point of the first quadrant; indifference curves corresponding to increasing levels of utility are farther from the origin. Indifference curves cannot intersect because this would imply that the same combination of q_1 and q_2 has different utilities.

Note that $U = f(q_1, q_2)$ can be represented by a utility surface in three dimensions. Each indifference curve represents the intersection of that surface and a plane parallel to the $q_1 q_2$ -plane at a certain value of U . The points on an indifference curve represent the various combinations of quantities of the commodities that have the same utility for the consumer and among which, therefore, he is indifferent.

The total differential of the utility function $U = f(q_1, q_2)$ is

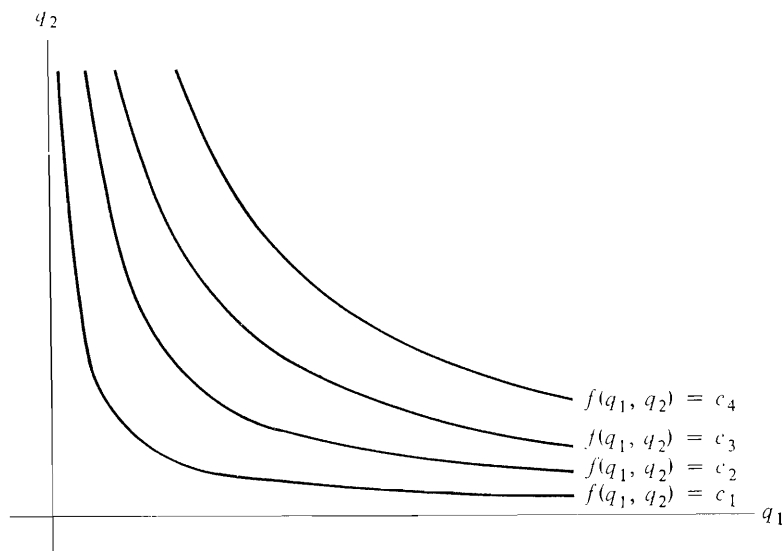
$$dU = f_1 dq_1 + f_2 dq_2$$

where f_1 and f_2 are the partial derivatives of U with respect to q_1 and q_2 , respectively. Since the change in utility along any indifference curve is zero by definition, $dU = 0$ and

$$f_1 dq_1 + f_2 dq_2 = 0$$

$$-\frac{dq_2}{dq_1} = \frac{f_1}{f_2}$$

Figure 3.5



The slope of an indifference curve dq_2/dq_1 is the rate at which a consumer would be willing to substitute Q_1 for Q_2 or Q_2 for Q_1 to maintain a given level of utility. The negative of the slope— $-dq_2/dq_1$ is the *rate of commodity substitution or marginal rate of substitution* of Q_1 for Q_2 or Q_2 for Q_1 and equals the ratio of the partial derivatives of the utility function. Note that the rate of commodity substitution at a point on an indifference curve is the same for movements in either direction, so the verbal definition can be in terms of substituting Q_1 for Q_2 or Q_2 for Q_1 .

The partial derivatives f_1 and f_2 are said to be marginal utilities of the commodities Q_1 and Q_2 . The magnitudes of the marginal utilities are without meaning when the utility function is ordinal, but their signs and ratios are meaningful. Note that if the marginal utilities are positive, an increase in the quantity consumed of one commodity with no change in the quantity consumed of the other commodity will move the consumer to a higher indifference curve.

PROBLEMS

- For each of the following production functions, find the marginal productivities. Output is denoted by z and the inputs are denoted by x and y .

a. $z = 25 - \frac{1}{x} - \frac{1}{y}$ at $x = 1, y = 1$

b. $z = 5xy - 2x^2 - 2y^2$ at $x = 1, y = 1$

c. $16z^2 - z - 80 + 4(x - 5)^2 + 2(y - 4)^2 = 0$

d. $6z^3 - z^2 - 6x - 24y + x^2 + 4y^2 + 50 = 0$

- The Cobb-Douglas production function for the economy as a whole is given by

$$z = ax^by^c$$

where z is total product, x is quantity of labor, y is quantity of capital, and a , b , and c are constants. It is frequently assumed that $b + c = 1$. Is this function homogeneous and, if so, of what degree?

3. For each of the following production functions, determine the degree of homogeneity and the nature of the returns to scale. Output is denoted by z and the inputs are denoted by x and y .

a. $z = 3x^3 + 5xy^2 + y^3$

c. $z = 25y^6 - x^2y^4$

b. $z = \frac{14}{x} - \frac{20}{y}$

d. $z = \frac{3}{x^2} + \frac{25}{xy} + \frac{6}{y^2}$

4. In addition to the property known as Euler's theorem, linear homogeneous functions have the following properties: If $z = f(x, y)$ is a linear homogeneous function, then

$$\left\{ \begin{array}{l} \frac{z}{x} = g_1\left(\frac{y}{x}\right) \quad \text{that is, } \frac{z}{x} \text{ is a function of } \frac{y}{x} \\ \frac{z}{y} = g_2\left(\frac{x}{y}\right) \quad \text{that is, } \frac{z}{y} \text{ is a function of } \frac{x}{y} \end{array} \right.$$

$$\left\{ \begin{array}{l} \frac{\partial z}{\partial x} = h_1\left(\frac{y}{x}\right) \quad \text{that is, } \frac{\partial z}{\partial x} \text{ is a function of } \frac{y}{x} \\ \frac{\partial z}{\partial y} = h_2\left(\frac{x}{y}\right) \quad \text{that is, } \frac{\partial z}{\partial y} \text{ is a function of } \frac{x}{y} \end{array} \right.$$

$$\left\{ \begin{array}{l} \frac{\partial^2 z}{\partial x^2} = -\frac{y}{x} \cdot \frac{\partial^2 z}{\partial x \partial y} \\ \frac{\partial^2 z}{\partial y^2} = -\frac{x}{y} \cdot \frac{\partial^2 z}{\partial x \partial y} \end{array} \right.$$

$$\left\{ \begin{array}{l} \frac{\partial^2 z}{\partial x^2} = -\frac{y}{x} \cdot \frac{\partial^2 z}{\partial x \partial y} \\ \frac{\partial^2 z}{\partial y^2} = -\frac{x}{y} \cdot \frac{\partial^2 z}{\partial x \partial y} \end{array} \right.$$

$$\left\{ \begin{array}{l} \frac{\partial^2 z}{\partial x^2} = -\frac{y}{x} \cdot \frac{\partial^2 z}{\partial x \partial y} \\ \frac{\partial^2 z}{\partial y^2} = -\frac{x}{y} \cdot \frac{\partial^2 z}{\partial x \partial y} \end{array} \right.$$

Determine whether each of the following functions is homogeneous; for homogeneous functions, determine the degree and demonstrate Euler's theorem; for linear homogeneous functions, also demonstrate the other three properties given above.

a. $z = 3x^2 + 4xy + 15y^3$

i. $z = 4x^2e^y + 3y^2e^x$

b. $z = 4x^3 + x^2y - 3xy^2 - 7y^3$

j. $z = 3x^2 \ln a^{1/y^2} - 5y^3 \ln b^{1/y^3}$

c. $z = 3e^x + 3e^y$

k. $z = \frac{x^2 + xy}{3y}$

d. $z = 6 \ln 3^{-2x} - \ln 4^{-5y}$

l. $z = 3x^2y + 4xy^2 + y^3 + 10$

e. $z = \frac{x^4 + 3x^2y^2 + xy^3 + 6y^4}{x^3}$

m. $z = \frac{xy}{x + y}$

f. $z = 3x \ln 5^{x/y} - 9y \ln 8^{y/x}$

n. $z = xy - \ln xy$

g. $z = \frac{x^2y + y^3}{x^6}$

h. $z = \frac{x^2 + 3xy + y^2}{y^2}$

5. If a consumer's utility function is given by $U = q_1 q_2^2$ and the consumer purchases 4 units of Q_1 and 5 units of Q_2 , a. what quantity of Q_1 must he purchase to maintain the same level of utility if his purchase of Q_2 increases to 6 units? b. What quantity of Q_2 must he purchase to maintain the same level of utility if his purchase of Q_1 increases to 6 units? c. What quantity of Q_1 must he purchase to maintain the same level of utility if his purchase of Q_2 decreases to 4 units? d. What quantity of Q_2 must he purchase to maintain the same level of utility if his purchase of Q_1 decreases to 2 units?

6. Find the marginal utilities for each of the following utility functions. Utility is denoted by U and the amounts of the two commodities consumed are denoted by q_1 and q_2 , respectively.
- $U = q_1^2 q_2$
 - $U = q_1 q_2 + q_1^2$

Answers to Odd-Numbered Problems

- $\frac{\partial z}{\partial x} = 1, \frac{\partial z}{\partial y} = 1$
 - $\frac{\partial z}{\partial x} = 1, \frac{\partial z}{\partial y} = 1$
 - $\frac{\partial z}{\partial x} = -\frac{8(x-5)}{32z-1}$
 $\frac{\partial z}{\partial y} = -\frac{4(y-4)}{32z-1}$
 - $\frac{\partial z}{\partial x} = -\frac{x-3}{z(9z-1)}$
 $\frac{\partial z}{\partial y} = -\frac{4(y-3)}{z(9z-1)}$
- degree 3, increasing returns to scale
 - degree -1, decreasing returns to scale
 - degree 6, increasing returns to scale
 - degree -2, decreasing returns to scale
- $q_1 = 25/9$
 - $q_2 = \frac{5\sqrt{6}}{3}$
 - $q_1 = 25/4$
 - $q_2 = 5\sqrt{2}$

3.4 Maxima and Minima of Functions of Two Variables

Just as the first and second derivatives of a function of one variable can be used to determine its maxima and minima, the first-order and second-order partial derivatives of a function of two variables can be used to determine its maxima and minima. The procedure becomes more complicated as the number of variables increases, but the definitions and methods are essentially the same. The discussion below concerns functions of two independent variables; maxima and minima of functions of n variables are discussed in Chapter 8.

A function $f(x, y)$ of two independent variables is said to have a local maximum value (or minimum value) for $x = a, y = b$, if $f(a, b)$ is greater (or less) than $f(x, y)$ for all values of x and y close to $x = a, y = b$. If $f(x, y)$ has a maximum (or minimum) value at $x = a, y = b$, it follows that the function $f(x, b)$ has a maximum (or minimum) at $x = a$ and the function $f(a, y)$ has a maximum (or minimum) at $y = b$. Thus, for the function $f(x, y)$ to have a maximum (or minimum) at $x = a, y = b$, it is necessary that

$$\left. \frac{\partial}{\partial x} f(x, y) \right|_{x=a, y=b} = 0$$

$$\left. \frac{\partial}{\partial y} f(x, y) \right|_{x=a, y=b} = 0$$

These two conditions are used to determine the critical points; the procedure below is then used to determine whether the critical points are in fact local maxima or minima.

$$\text{If } \begin{cases} \left. \frac{\partial}{\partial x} f(x, y) \right|_{x=a, y=b} = 0 \\ \left. \frac{\partial}{\partial y} f(x, y) \right|_{x=a, y=b} = 0 \end{cases}$$

and

$$\Delta = \left(\left. \frac{\partial^2}{\partial x^2} f(x, y) \right|_{x=a, y=b} \right) \left(\left. \frac{\partial^2}{\partial y^2} f(x, y) \right|_{x=a, y=b} \right) - \left(\left. \frac{\partial^2}{\partial x \partial y} f(x, y) \right|_{x=a, y=b} \right)^2$$

$$\text{then } \begin{cases} \Delta > 0 \Rightarrow \begin{cases} \text{maximum at } x = a, y = b, \text{ if } \left. \frac{\partial^2}{\partial x^2} f(x, y) \right|_{x=a, y=b} < 0 \text{ and } \left. \frac{\partial^2}{\partial y^2} f(x, y) < 0 \\ \text{minimum at } x = a, y = b, \text{ if } \left. \frac{\partial^2}{\partial x^2} f(x, y) > 0 \text{ and } \left. \frac{\partial^2}{\partial y^2} f(x, y) > 0 \end{cases} \\ \Delta < 0 \Rightarrow \text{neither maximum nor minimum at } x = a, y = b, \text{ but a saddle point at } x = a, y = b \\ \Delta = 0 \Rightarrow \text{test fails, function must be investigated near } x = a, y = b \end{cases}$$

These conditions can be stated more concisely as follows:

necessary condition for a critical point: $f_x = 0, f_y = 0$

determination of local maxima or minima:

$$f_{xx} f_{yy} - f_{xy}^2 > 0 \Rightarrow \begin{cases} \text{maximum if } f_{xx} < 0, f_{yy} < 0 \\ \text{minimum if } f_{xx} > 0, f_{yy} > 0 \end{cases}$$

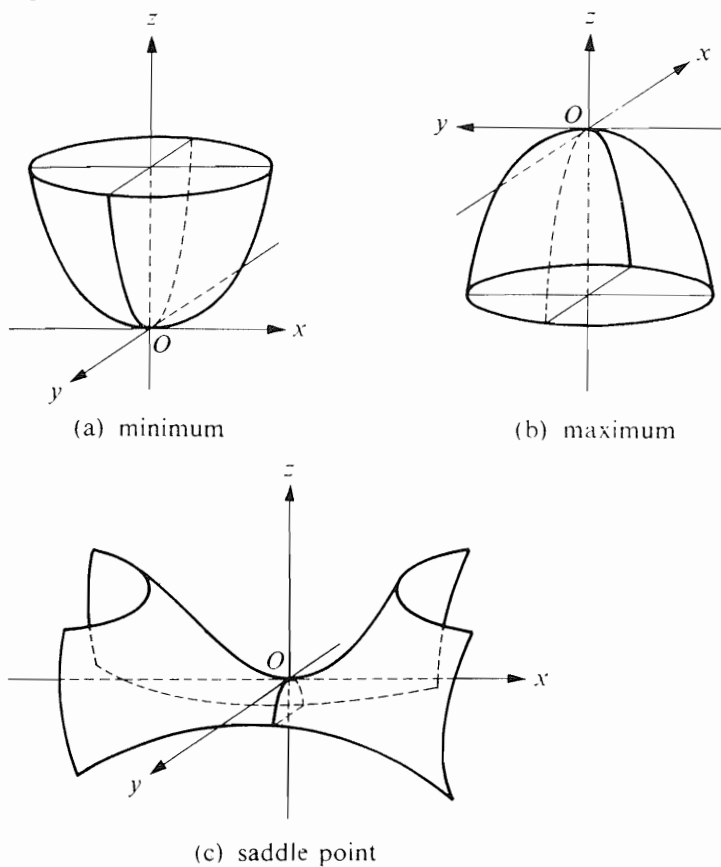
$$f_{xx} f_{yy} - f_{xy}^2 < 0 \Rightarrow \text{saddle point}$$

$$f_{xx} f_{yy} - f_{xy}^2 = 0 \Rightarrow \text{test fails and the function must be investigated near the critical point where the partial derivatives are evaluated for the critical point}$$

Note that this procedure is used to identify local maxima and minima. In order to determine global maxima or minima it is necessary not only to compare the values of the local maxima or minima, but also to determine the values of the function at the limits of its range, if its range is limited. The procedure for determining global maxima and minima for functions of two variables is thus analogous to the procedure for functions of one variable.

Examples of functions having at the origin (a) a minimum, (b) a maximum, and (c) a saddle point are illustrated in Figure 3.6. Note that at a saddle point a function has a minimum with respect to one variable and a maximum with respect to the other variable. For example, in Figure 3.6(c) the origin looks like a minimum to a person traveling along the surface in the yz -plane, but to a person traveling in the xz -plane the origin looks like a maximum.

Figure 3.6

**EXAMPLE**

Examine the function

$$f(x, y) = x^2 + xy + y^2 - 3x + 2$$

for maximum and minimum values.

$$\frac{\partial f}{\partial x} = 2x + y - 3$$

$$\frac{\partial f}{\partial y} = x + 2y$$

$$2x + y = 3$$

$$x + 2y = 0$$

$$3y = -3$$

$$y = -1, x = 2 \quad \text{so } (2, -1) \text{ may be a maximum or minimum}$$

$$\frac{\partial^2 f}{\partial x^2} = 2$$

$$\frac{\partial^2 f}{\partial y^2} = 2$$

$$\frac{\partial^2 f}{\partial x \partial y} = 1$$

$$\Delta = 4 - 1 > 0$$

$$\frac{\partial^2 f}{\partial x^2} = 2 > 0 \quad \frac{\partial^2 f}{\partial y^2} = 2 > 0$$

Thus $f(x, y)$ has a minimum value at $(2, -1)$.

EXAMPLE

Examine the function

$$f(x, y) = xy - \ln(x^2 + y^2)$$

for maximum and minimum values.

$$\frac{\partial f}{\partial x} = y - \frac{2x}{x^2 + y^2}$$

$$\frac{\partial f}{\partial y} = x - \frac{2y}{x^2 + y^2}$$

$$x^2 y + y^3 - 2x = 0$$

$$x^3 + xy^2 - 2y = 0$$

$$-2x^2 + 2y^2 = 0$$

$$x^2 = y^2$$

Substituting for x^2 in the first equation and for y^2 in the second equation, $y^3 = x$ and $x^3 = y$, so $(1, 1)$ and $(-1, -1)$ may be a maximum or minimum.

$$\frac{\partial^2 f}{\partial x^2} = -\frac{2(x^2 + y^2) - 2x(2x)}{(x^2 + y^2)^2} = -\frac{2(y^2 - x^2)}{(x^2 + y^2)^2}$$

$$\frac{\partial^2 f}{\partial y^2} = -\frac{2(x^2 + y^2) - 2y(2y)}{(x^2 + y^2)^2} = -\frac{2(x^2 - y^2)}{(x^2 + y^2)^2}$$

$$\frac{\partial f}{\partial x \partial y} = 1 + \frac{4xy}{(x^2 + y^2)^2}$$

$$\left. \frac{\partial^2 f}{\partial x^2} \right|_{x=y=1} = \left. \frac{\partial^2 f}{\partial x^2} \right|_{x=y=-1} = 0$$

$$\left. \frac{\partial^2 f}{\partial y^2} \right|_{x=y=1} = \left. \frac{\partial^2 f}{\partial y^2} \right|_{x=y=-1} = 0$$

$$\left. \frac{\partial^2 f}{\partial x \partial y} \right|_{x=y=1} = \left. \frac{\partial^2 f}{\partial x \partial y} \right|_{x=y=-1} = 2$$

$$\Delta = 0 - 2 < 0$$

Thus $f(x, y)$ has no maximum or minimum, but does have saddle points at $(1, 1)$ and $(-1, -1)$.

EXAMPLE

Examine the function

$$f(x, y) = 1 + x^2 - y^2$$

for maximum and minimum values.

$$\frac{\partial f}{\partial x} = 2x$$

$$\frac{\partial f}{\partial y} = -2y$$

$$\left. \begin{array}{l} 2x = 0 \\ -2y = 0 \end{array} \right\} \begin{array}{l} \text{if } x = 0 \\ \text{if } y = 0 \end{array} \left. \vphantom{\begin{array}{l} 2x = 0 \\ -2y = 0 \end{array}} \right\} \text{so } (0, 0) \text{ may be a maximum or minimum}$$

$$\frac{\partial^2 f}{\partial x^2} = 2$$

$$\frac{\partial^2 f}{\partial y^2} = -2$$

$$\frac{\partial^2 f}{\partial x \partial y} = 0$$

$$\Delta = -4 - 0 < 0$$

Thus $f(x, y)$ has neither a maximum nor a minimum at $(0, 0)$, but does have a saddle point at $(0, 0)$.

EXAMPLE

Examine the function

$$f(x, y) = \sin x + \sin y + \sin(x + y)$$

for maximum and minimum values.

$$\frac{\partial f}{\partial x} = \cos x + \cos(x + y)$$

$$\frac{\partial f}{\partial y} = \cos y + \cos(x + y)$$

If $\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0$, $\cos x = \cos y = -\cos(x + y)$, so $(\pi/3, \pi/3)$ may be a maximum or minimum.

$$\frac{\partial^2 f}{\partial x^2} = -\sin x - \sin(x + y)$$

$$\frac{\partial^2 f}{\partial y^2} = -\sin y - \sin(x + y)$$

$$\frac{\partial^2 f}{\partial x \partial y} = -\sin(x + y)$$

$$\left. \frac{\partial^2 f}{\partial x^2} \right|_{x=y=\pi/3} = -\frac{\sqrt{3}}{2} - \frac{\sqrt{3}}{2} = -\sqrt{3}$$

$$\left. \frac{\partial^2 f}{\partial y^2} \right|_{x=y=\pi/3} = -\frac{\sqrt{3}}{2} - \frac{\sqrt{3}}{2} = -\sqrt{3}$$

$$\left. \frac{\partial^2 f}{\partial x \partial y} \right|_{x=y=\pi/3} = -\frac{\sqrt{3}}{2}$$

$$\Delta = (-\sqrt{3})(-\sqrt{3}) - \left(-\frac{\sqrt{3}}{2}\right)^2$$

$$= 3 - \frac{3}{4} = \frac{9}{4} > 0$$

$$\frac{\partial^2 f}{\partial x^2} < 0 \quad \text{and} \quad \frac{\partial^2 f}{\partial y^2} < 0$$

Thus $f(x, y)$ has a maximum value for $x = y = \pi/3$.

EXAMPLE

Examine the function

$$f(x, y) = 25 + (x - y)^4 + (y - 1)^4$$

for maximum and minimum values.

$$\frac{\partial f}{\partial x} = 4(x - y)^3$$

$$\frac{\partial f}{\partial y} = -4(x - y)^3 + 4(y - 1)^3$$

$$\left. \begin{array}{ll} (x - y)^3 = 0 & \text{if } x = y \\ (x - y)^3 = (y - 1)^3 & \text{if } x = y = 1 \end{array} \right\} \text{ so } (1, 1) \text{ may be a maximum or minimum}$$

$$\frac{\partial^2 f}{\partial x^2} = 12(x - y)^2$$

$$\frac{\partial^2 f}{\partial y^2} = 12(x - y)^2 + 12(y - 1)^2$$

$$\frac{\partial^2 f}{\partial x \partial y} = -12(x - y)^2$$

$$\left. \frac{\partial^2 f}{\partial x^2} \right|_{x=y=1} = 0$$

$$\left. \frac{\partial^2 f}{\partial y^2} \right|_{x=y=1} = 0$$

$$\left. \frac{\partial^2 f}{\partial x \partial y} \right|_{x=y=1} = 0$$

so $\Delta = 0$ and $f(x, y)$ must be examined further for maximum or minimum values. Suppose h and k are arbitrarily small positive or negative numbers; then

$$\begin{aligned} f(1 + h, 1 + k) - f(1, 1) &= 25 + [(1 + h) - (1 + k)]^4 + (1 + k - 1)^4 - 25 \\ &= (h - k)^4 + k^4 \end{aligned}$$

But $(h - k)^4 + k^4 > 0$ for all h and k , so $f(1 + h, 1 + k) > f(1, 1)$ for all h and k , and thus there is a minimum at $(1, 1)$.

EXAMPLE

Examine the function

$$f(x, y) = x^2 - 6xy + 9y^2 + 3x - 10$$

for maximum and minimum values.

$$\frac{\partial f}{\partial x} = 2x - 6y + 3$$

$$\frac{\partial f}{\partial y} = -6x + 18y$$

$$2x - 6y = -3$$

$$-2x + 6y = 0$$

no simultaneous solution, so no maximum or minimum or saddle point

EXAMPLE

If the demand functions are

$$p = 36 - 3x \quad \text{and} \quad q = 40 - 5y$$

and the joint-cost function is

$$C = x^2 + 2xy + 3y^2$$

determine the quantities and prices that maximize profit for the monopolist and find the maximum profit.

$$\begin{aligned} P &= px + qy - C \\ &= 36x - 3x^2 + 40y - 5y^2 - x^2 - 2xy - 3y^2 \\ &= -4x^2 - 8y^2 - 2xy + 36x + 40y \end{aligned}$$

$$\frac{\partial P}{\partial x} = -8x - 2y + 36$$

$$\frac{\partial P}{\partial y} = -16y - 2x + 40$$

$$4x + y = 18$$

$$x + 8y = 20$$

$$31y = 62$$

$$\left. \begin{array}{l} y = 2 \\ x = 4 \end{array} \right\} \text{ so } P \text{ may have a maximum or minimum at } (4, 2)$$

$$\frac{\partial^2 P}{\partial x^2} = -8$$

$$\frac{\partial^2 P}{\partial y^2} = -16$$

$$\frac{\partial^2 P}{\partial x \partial y} = -2$$

$$\Delta = 128 - (-2)^2 = 124 > 0$$

$$\frac{\partial^2 P}{\partial x^2} < 0 \quad \text{and} \quad \frac{\partial^2 P}{\partial y^2} < 0,$$

so P has a maximum value for $x = 4$, $y = 2$. If $x = 4$, $y = 2$, then $p = 24$, $q = 30$, and $P_{\max} = 112$.

EXAMPLE

Suppose the production function is

$$16z = 65 - 2(x - 5)^2 - 4(y - 4)^2$$

The unit prices of the inputs x and y (under pure competition) are 8 and 4, respectively, and the unit price of the output is 32; determine the maximum profit.

$$P = 32z - 8x - 4y$$

$$= 130 - 4(x - 5)^2 - 8(y - 4)^2 - 8x - 4y$$

$$\frac{\partial P}{\partial x} = -8(x - 5) - 8 = -8x + 32$$

$$\frac{\partial P}{\partial y} = -16(y - 4) - 4 = -16y + 60$$

$$\left. \begin{array}{l} 8x = 32 \\ x = 4 \\ 16y = 60 \\ y = \frac{15}{4} \end{array} \right\} \text{ so } P \text{ may have a maximum or minimum at } (4, \frac{15}{4})$$

$$\frac{\partial^2 P}{\partial x^2} = -8$$

$$\frac{\partial^2 P}{\partial y^2} = -16$$

$$\frac{\partial^2 P}{\partial x \partial y} = 0$$

$$\Delta = (-8)(-16) - 0 = 128 > 0$$

$$\frac{\partial^2 P}{\partial x^2} < 0 \quad \text{and} \quad \frac{\partial^2 P}{\partial y^2} < 0,$$

so P has a maximum value for $x = 4$, $y = \frac{15}{4}$ and $P_{\max} = 78\frac{1}{2}$.

EXAMPLE

In simple bivariate regression analysis, the model involves the assumption of a linear relationship between a predicted variable y and an explanatory variable x and a disturbance term u . For a sample of n observations of x and y ,

$$y_i = \alpha + \beta x_i + u_i \quad i = 1, 2, \dots, n$$

and the problem is to obtain estimates of α and β . These estimates are customarily obtained using the least squares principle. Let $\hat{\alpha}$ and $\hat{\beta}$ denote estimates of α and β . Then $y_i = \hat{\alpha} + \hat{\beta}x_i + e_i$ where e_i denotes the i th residual $y_i - \hat{\alpha} - \hat{\beta}x_i$. The least squares principle minimizes the sum of the squared residuals

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat{\alpha} - \hat{\beta}x_i)^2$$

Differentiating partially with respect to $\hat{\alpha}$ and $\hat{\beta}$,

$$\frac{\partial}{\partial \hat{\alpha}} \left(\sum_{i=1}^n e_i^2 \right) = -2 \sum_{i=1}^n (y_i - \hat{\alpha} - \hat{\beta}x_i)$$

$$\frac{\partial}{\partial \hat{\beta}} \left(\sum_{i=1}^n e_i^2 \right) = -2 \sum_{i=1}^n x_i (y_i - \hat{\alpha} - \hat{\beta}x_i)$$

Equating the partial derivatives to zero,

$$\sum_{i=1}^n y_i - n\hat{\alpha} - \hat{\beta} \sum_{i=1}^n x_i = 0$$

$$\sum_{i=1}^n x_i y_i - \hat{\alpha} \sum_{i=1}^n x_i - \hat{\beta} \sum_{i=1}^n x_i^2 = 0$$

Multiplying the first equation by $\sum_{i=1}^n x_i$ and the second equation by n and solving the equations simultaneously,

$$\hat{\beta} = \frac{n \sum_{i=1}^n x_i y_i - \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right)}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}$$

$$\hat{\alpha} = \frac{\sum_{i=1}^n y_i - \hat{\beta} \sum_{i=1}^n x_i}{n}$$

are the least squares estimates for the regression coefficients α and β .

To show that this solution minimizes $\sum_{i=1}^n e_i^2$, the second order condition for a minimum must be considered.

$$\frac{\partial^2}{\partial \hat{\alpha}^2} \left(\sum_{i=1}^n e_i^2 \right) = 2n$$

$$\frac{\partial^2}{\partial \hat{\beta}^2} \left(\sum_{i=1}^n e_i^2 \right) = 2 \sum_{i=1}^n x_i^2$$

$$\frac{\partial^2}{\partial \hat{\alpha} \partial \hat{\beta}} \left(\sum_{i=1}^n e_i^2 \right) = 2 \sum_{i=1}^n x_i$$

$$\Delta = (2n)(2 \sum_{i=1}^n x_i^2) - (2 \sum_{i=1}^n x_i)^2 > 0, \text{ since } n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 > 0 \text{ and}$$

$$\frac{\partial^2}{\partial \hat{\alpha}^2} \left(\sum_{i=1}^n e_i^2 \right) > 0 \quad \frac{\partial^2}{\partial \hat{\beta}^2} \left(\sum_{i=1}^n e_i^2 \right) > 0$$

Thus $\hat{\alpha}$, $\hat{\beta}$ minimize $\sum_{i=1}^n e_i^2$ and are the least squares estimators of α and β .

EXAMPLE

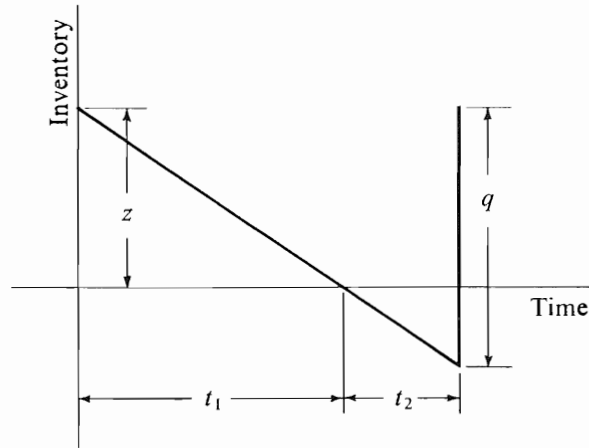
A manufacturer has a demand of D units per period T . The storage cost per period per unit is c_2 dollars, the shortage cost is c_3 dollars per unit short per period, and it costs c_1 dollars each time a production run is started (setup cost). Assuming that production orders are filled with no delay and demand is at a uniform rate, determine the frequency with which production should be scheduled, and the quantity which should be produced in each run in order to minimize the total per period inventory cost K . In Figure 3.7, t = interval between production orders, q = production quantity, z = planned inventory at completion of a production run, t_1 = time (during cycle) when stock is on hand, and t_2 = time (during cycle) when no stock is on hand. Note that shortages are assumed to be backlogged and not lost.

$$\text{Note that } \frac{t_1}{t} = \frac{z}{q}, \quad \frac{t_2}{t} = \frac{q-z}{q}, \text{ and } \frac{t}{T} = \frac{q}{D}.$$

setup cost: c_1

$$\text{carrying cost per run: } \frac{c_2 t_1 z}{2} = \frac{c_2 t z^2}{2q} = \frac{c_2 T z^2}{2D}$$

Figure 3.7



$$\text{shortage cost per run: } \frac{c_3 t_2 (q - z)}{2} = \frac{c_3 t (q - z)^2}{2q} = \frac{c_3 T (q - z)^2}{2D}$$

since the number of runs per period is $\frac{D}{q}$, the total inventory cost per period is

$$C = \frac{c_1 D}{q} + \frac{c_2 T z^2}{2q} + \frac{c_3 T (q - z)^2}{2q}$$

and, since C is a function of both q and z , partial derivatives must be obtained and set equal to zero to determine a minimum

$$\frac{\partial C}{\partial z} = \frac{c_2 T z}{q} - \frac{c_3 T (q - z)}{q}$$

$$\text{If } \frac{\partial C}{\partial z} = 0,$$

$$c_2 T z = c_3 T (q - z)$$

$$z = \frac{c_3 q}{c_2 + c_3}$$

$$\begin{aligned} \frac{\partial C}{\partial q} &= -\frac{c_1 D}{q^2} - \frac{c_2 T z^2}{2q^2} + \frac{c_3 T}{2} - \frac{c_3 T z^2}{2q^2} \\ &= -\frac{c_1 D}{q^2} - \frac{c_2 T z^2}{2q^2} + \frac{c_3 T (q^2 - z^2)}{2q^2} \end{aligned}$$

$$\text{If } \frac{\partial C}{\partial q} = 0,$$

$$c_3 T (q^2 - z^2) = 2c_1 D + c_2 T z^2$$

$$q^2 = \frac{2c_1 D + c_2 T z^2 + c_3 T z^2}{c_3 T}$$

$$= \frac{2c_1 D}{c_3 T} + \frac{c_2 + c_3}{c_3} z^2$$

$$\text{Substituting } z = \frac{c_3 q}{c_2 + c_3},$$

$$c_3 T (q^2 - z^2) = \frac{2c_1 D}{c_3 T} + \frac{c_3}{c_2 + c_3} q^2$$

thus

$$q^2 = \frac{2c_1 D}{\left(1 - \frac{c_3}{c_2 + c_3}\right) c_3 T}$$

$$q = \sqrt{\frac{2c_1(c_2 + c_3)D}{c_2 c_3 T}}$$

$$z = \frac{c_3}{c_2 + c_3} q = \sqrt{\frac{2c_1 c_3 D}{c_2(c_2 + c_3)T}}$$

$$t = \frac{Tq}{D} = \sqrt{\frac{2c_1(c_2 + c_3)T}{c_2 c_3 D}}$$

The manufacturer should produce q units at intervals of time t , where q and t are determined by the formulas above.

PROBLEMS

Determine the maxima, minima, and saddle points (if any) for each of the following functions.

1. $g(x, y) = 2x^2 - 2xy + y^2 + 5x - 3y$
2. $h(x, y) = 3 + 2x + 2y - 2x^2 - 2xy - y^2$
3. $z(x, y) = xy + x - y$
4. $f(x, y) = x^2 + xy + y^2 - 6x + 2$
5. $u(x, y) = 4x + 2y - x^2 + xy - y^2$
6. $v(x, y) = x^2 + y^2 - 2x + 4y + 6$
7. $z(x, y) = x^3 - 3bxy + y^3$
8. $w(x, y) = x^2 - 2xy + 2y^2 - 2x + 2y + 1$
9. $g(x, y) = x^2 - y^2 - 2x + 4y + 6$
10. $h(x, y) = x^2 + 2xy$
11. $z(x, y) = xy - 2y^2$
12. $f(x, y) = axy$

Find the quantities and prices that maximize profit and the maximum profit for each of the following sets of demand functions and joint-cost functions for two commodities. Quantities are denoted by x and y and the corresponding prices are denoted by p and q , respectively; total cost is denoted by C .

- | | |
|--|--|
| 13. $x = 1 - p + 2q$
$y = 11 + p - 3q$
$C = 4x + y$ | 17. $p = 26 - x$
$q = 40 - 4y$
$C = x^2 + 2xy + y^2$ |
| 14. $x = 11 - 2p - 2q$
$y = 16 - 2p - 3q$
$C = 3x + y$ | 18. $p = 35 - 2x^2$
$q = 20 - y$
$C = 16 - 2x^3 + xy + 30x + 12y + \frac{1}{2}x^2$ |
| 15. $p = 40 - 2x^2$
$q = 12 - 3y$
$C = 8 + 4x + 3y$ | 19. $p = 40 - 5x$
$q = 30 - 3y$
$C = x^2 + 2xy + 3y^2$ |
| 16. $p = 16 - x^2$
$q = 9 - y^2$
$C = x^2 + 3y^2$ | 20. $p = 28 - 3x^2$
$q = 56 - y^2$
$C = 2x^2 + y^2$ |
21. Determine the maximum profit if the production function is $z = 20 - x^2 + 10x - 2y^2 + 5y$, the prices of the inputs x and y are 2 and 1, respectively, and the price of the output z is 5.
 22. Determine the maximum profit if the production function is $z = 10 - 2x^2 + xy - y^2 + 5y$, the prices of the inputs x and y are each 3, and the price of the output z is 6.

23. The Portland Company has a contract to furnish 1350 cement mixers annually. The yearly storage cost is \$40 per cement mixer, the shortage cost is \$50 per cement mixer short per year, and it costs \$150 to start a production run. If production orders are filled without delay and demand is at a constant rate, determine the frequency with which production should be scheduled and the quantity which should be produced in each run in order to minimize the total average annual cost.
24. The Filaway Company has a contract to supply 600 filing cabinets at a uniform rate during a 9-month period. The storage cost during this period is \$30 per cabinet, the shortage cost is \$60 per cabinet, and it costs \$20 to start a production run. If production occurs at a constant rate of 2400 cabinets per 9-month period, determine the frequency with which production should be scheduled and the quantity which should be produced in each run in order to minimize the total average annual cost.

Answers to Odd-Numbered Problems

- | | |
|--|--|
| 1. minimum $(-1, \frac{1}{2})$ | 17. $p = 21$ |
| 3. saddle point $(1, -1)$ | $q = 28$ |
| 5. maximum $(\frac{10}{3}, \frac{8}{3})$ | $x = 5$ |
| 7. minimum (b, b) | $y = 3$ |
| 9. saddle point $(1, 2)$ | $P_{\max} = 125$ |
| 11. saddle point $(0, 0)$ | 19. $p = 25$ |
| 13. $p = 14$ | $q = 24$ |
| $q = 8$ | $x = 3$ |
| $x = 3$ | $y = 2$ |
| $y = 1$ | $P_{\max} = 90$ |
| $P_{\max} = 37$ | 21. $P_{\max} = 229\frac{2}{3}$ for $x = \frac{25}{5}$ and $y = \frac{6}{5}$ |
| 15. $p = 28$ | 23. Schedule production 10 times a year, |
| $q = \frac{15}{2}$ | producing 135 cement mixers each run. |
| $x = \sqrt{6}$ | |
| $y = \frac{3}{2}$ | |
| $P_{\max} = 24\sqrt{6} - \frac{5}{4}$ | |
| (approximately 57.53) | |

3.5 Maxima and Minima Subject to Constraints

In many practical applications of maximization and minimization, the problem is to maximize or minimize a given function subject to certain side conditions or constraints on the variables involved. These constraints may be stated as equalities or inequalities.

For example, if a manufacturer produces two outputs, he may want to minimize the joint cost, while producing a specified minimum total amount; a company may want to maximize sales resulting from use of two advertising media, while keeping total advertising costs within a specified budget constraint; an engineer may want to minimize frequency of breakdowns, while keeping total cost of two types of preventive repairs from exceeding a specified amount; a consumer may want to maximize utility derived from the consumption of commodities, subject to his budget constraint; a chemist may want to minimize total sedimentation in a two-stage process, while keeping the temperatures in the two stages equal.

In some cases equations obtained from the constraints can be substituted into the function to be maximized or minimized; the problem is thus reduced to one involving unconstrained maxima or minima and can be solved by the methods of the preceding section. However, this procedure is not always feasible, particularly if

the function to be maximized or minimized involves more than two variables and several constraints. The methods discussed in this section are applicable in all cases and can be generalized for use with any number of variables and constraints.

The method of Lagrange multipliers is used in obtaining maxima or minima of functions subject to equality constraints, and this method is easily generalized for the case of one inequality constraint. Maxima or minima subject to multiple equality and inequality constraints are obtained using the Kuhn-Tucker conditions.

LAGRANGE MULTIPLIERS

The most widely used method of obtaining maxima or minima of functions subject to equality constraints is *Lagrange multipliers*. Suppose $f(x, y)$ is to be maximized or minimized subject to the constraint $g(x, y) = 0$. Form the *objective function*

$$F(x, y, \lambda) = f(x, y) - \lambda g(x, y)$$

where λ , the Lagrange multiplier, is an unknown. Differentiate $F(x, y, \lambda)$ partially with respect to x , with respect to y , and with respect to λ , and set the results equal to zero; then the three equations

$$\frac{\partial F}{\partial x} = \frac{\partial f}{\partial x} - \lambda \frac{\partial g}{\partial x} = 0$$

$$\frac{\partial F}{\partial y} = \frac{\partial f}{\partial y} - \lambda \frac{\partial g}{\partial y} = 0$$

$$\frac{\partial F}{\partial \lambda} = g(x, y) = 0$$

can be solved for the three unknowns x , y , and λ . Note that $\frac{\partial F}{\partial \lambda} = g(x, y) = 0$ is the constraint, so $F(x, y, \lambda)$ really need be differentiated partially only with respect to x and with respect to y . In many cases the values of the λ 's (Lagrange multipliers) are not of interest and are not found; for this reason they are sometimes referred to as "undetermined multipliers." The Lagrange multiplier may be preceded by a plus sign rather than by a minus sign and this form is used by some authors. The only resulting change in the solution is a change in the sign of λ .

The logic of the method of Lagrange multipliers is quite straightforward. The problem is to determine possible maximum or minimum values of a function that will satisfy a constraint. If it can be assumed that the points being considered satisfy the constraint, this problem reduces to the problem of finding an unconstrained maximum or minimum. Essentially, the method of Lagrange multipliers screens the possible maximum or minimum values to obtain only those values that satisfy the constraint.

Note that, if the constraint is satisfied, $g(x, y) = 0$ and $\lambda g(x, y) = 0$, regardless of the value of λ . In that case, the objective function becomes the unconstrained function $f(x, y)$. By considering λ to be an additional variable and setting the partial derivative with respect to λ equal to zero, the constraint $g(x, y) = 0$ is incorporated into the problem and the possible maximum or minimum values satisfy the constraint.

Solution of the three equations above provides the critical points of the constrained function. These critical points satisfy the constraint, but must be tested as maxima or minima of the function by a procedure similar to that for unconstrained maxima or minima.

For a critical point $x = a, y = b$,

$$\text{If } \begin{cases} \left. \frac{\partial F}{\partial x} \right|_{x=a, y=b} = 0 \\ \left. \frac{\partial F}{\partial y} \right|_{x=a, y=b} = 0 \end{cases}$$

$$\text{and } \Delta^* = \left(\left. \frac{\partial^2 F}{\partial x^2} \right|_{x=a, y=b} \right) \left(\left. \frac{\partial^2 F}{\partial y^2} \right|_{x=a, y=b} \right) - \left(\left. \frac{\partial^2 F}{\partial x \partial y} \right|_{x=a, y=b} \right)^2$$

$$\text{then } \begin{cases} \Delta^* > 0 \Rightarrow \begin{cases} \text{maximum at } x = a, y = b, \text{ if } \left. \frac{\partial^2 F}{\partial x^2} \right|_{x=a, y=b} < 0 \text{ and } \left. \frac{\partial^2 F}{\partial y^2} \right|_{x=a, y=b} < 0 \\ \text{minimum at } x = a, y = b, \text{ if } \left. \frac{\partial^2 F}{\partial x^2} \right|_{x=a, y=b} > 0 \text{ and } \left. \frac{\partial^2 F}{\partial y^2} \right|_{x=a, y=b} > 0 \end{cases} \\ \Delta^* \leq 0 \Rightarrow \text{test fails, function must be investigated near } x = a, y = b \end{cases}$$

Note that for unconstrained maxima and minima, if $\Delta < 0$, the critical point is neither a maximum nor a minimum; however, for constrained maxima and minima, if $\Delta < 0$, the critical point may in fact be a maximum or minimum. This corresponds to the fact that a point may be a maximum or minimum of the constrained function, although it is not a maximum or minimum of the unconstrained function.

These conditions can be stated more concisely as follows.

necessary condition for a critical point: $f_x = 0, f_y = 0$

determination of constrained maxima or minima:

$$f_{xx}f_{yy} - f_{xy}^2 > 0 \Rightarrow \begin{cases} \text{maximum if } f_{xx} < 0, f_{yy} < 0 \\ \text{minimum if } f_{xx} > 0, f_{yy} > 0 \end{cases}$$

$$f_{xx}f_{yy} - f_{xy}^2 \leq 0 \Rightarrow \begin{cases} \text{test fails and the function must be investigated near the critical point} \\ \text{where the partial derivatives are evaluated for the critical point} \end{cases}$$

NOTE: The method of Lagrange multipliers can be extended to a function of n variables, $f(x_1, x_2, \dots, x_n)$, subject to the k constraints $g_j(x_1, x_2, \dots, x_n) = 0, j = 1, 2, \dots, k$, where $k \leq n$. Then

$$F(x_1, \dots, x_n; \lambda_1, \dots, \lambda_k) = f(x_1, \dots, x_n) - \sum_{j=1}^k \lambda_j g_j(x_1, \dots, x_n)$$

and partial differentiation results in $n + k$ equations to be solved for $n + k$ unknowns. The method of testing critical points as maxima or minima can also be extended for the general case.

EXAMPLE

Find the maxima and minima (if any) of $f(x, y) = 5x^2 + 6y^2 - xy$ subject to the constraint $x + 2y = 24$.

$$F(x, y, \lambda) = 5x^2 + 6y^2 - xy - \lambda(x + 2y - 24)$$

$$\frac{\partial F}{\partial x} = 10x - y - \lambda$$

$$\frac{\partial F}{\partial y} = 12y - x - 2\lambda$$

$$\frac{\partial F}{\partial \lambda} = x + 2y - 24$$

Eliminating λ from the first two equations,

$$20x - 2y - 2\lambda = 0$$

$$-x + 12y - 2\lambda = 0$$

$$21x - 14y = 0$$

$$3x - 2y = 0$$

Solving simultaneously with the third equation,

$$x + 2y = 24$$

$$4x = 24$$

$$x = 6$$

Critical point: $(6, 9)$; $\lambda = 51$.

$$\frac{\partial^2 F}{\partial x^2} = 10$$

$$\frac{\partial^2 F}{\partial y^2} = 12$$

$$\frac{\partial^2 F}{\partial x \partial y} = -1$$

$$\Delta^* = (10)(12) - (-1)^2 = 119$$

$$\frac{\partial^2 F}{\partial x^2} > 0 \quad \text{and} \quad \frac{\partial^2 F}{\partial y^2} > 0$$

so $(6, 9)$ is a minimum.

$$\text{Alternative method: } f(x, y) = 5x^2 + 6y^2 - xy, \quad x + 2y = 24.$$

$$G(y) = 5(24 - 2y)^2 + 6y^2 - y(24 - 2y)$$

$$= 2880 - 480y + 20y^2 + 6y^2 - 24y + 2y^2$$

$$= 2880 - 504y + 28y^2$$

$$\frac{dG}{dy} = -504 + 56y$$

$$= 0 \quad \text{if } 56y = 504$$

$$y = 9$$

$$x = 6$$

$$\frac{d^2 G}{dy^2} = 56 > 0$$

so $(6, 9)$ is a minimum.

EXAMPLE

Find the maxima and minima (if any) of $f(x, y) = 12xy - 3y^2 - x^2$ subject to the constraint $x + y = 16$.

$$F(\lambda, x, y) = 12xy - 3y^2 - x^2 - \lambda(x + y - 16)$$

$$\frac{\partial F}{\partial x} = 12y - 2x - \lambda$$

$$\frac{\partial F}{\partial y} = 12x - 6y - \lambda$$

$$\frac{\partial F}{\partial \lambda} = -(x + y - 16)$$

Eliminating λ from the first two equations,

$$14x - 18y = 0$$

$$7x - 9y = 0$$

Solving simultaneously with the third equation,

$$7x + 7y = 112$$

$$16y = 112$$

$$y = 7$$

$$x = 9$$

Critical point: $(9, 7)$; $\lambda = 66$.

$$\frac{\partial^2 F}{\partial x^2} = -2$$

$$\frac{\partial^2 F}{\partial y^2} = -6$$

$$\frac{\partial^2 F}{\partial x \partial y} = 12$$

$$\Delta^* = (-2)(-6) - (12)^2 = -132 < 0$$

so the function must be investigated near $(9, 7)$.

$$(x + h) + (y + k) = 16$$

$$9 + h + 7 + k = 16$$

$$k = -h$$

$$\begin{aligned} f(9 + h, 7 - h) - f(9, 7) &= 12(9 + h)(7 - h) - 3(7 - h)^2 - (9 + h)^2 - 12(9)(7) - 3(7)^2 - 9^2 \\ &= -24h - 12h^2 + 42h - 3h^2 - 18h - h^2 \\ &= -16h^2 < 0 \end{aligned}$$

so maximum is at $(9, 7)$.

Alternative method: $f(x, y) = 12xy - 3y^2 - x^2$, $x + y = 16$.

$$\begin{aligned} G(y) &= 12y(16 - y) - 3y^2 - (16 - y)^2 \\ &= 192y - 12y^2 - 3y^2 - 256 + 32y - y^2 \\ &= -16y^2 + 224y - 256 \end{aligned}$$

$$\frac{dG}{dy} = -32y + 224$$

$$\frac{dG}{dy} = 0 \quad \text{if } 32y = 224$$

$$y = 7$$

$$x = 9$$

$$\frac{d^2G}{dy^2} = -32$$

so maximum is at (9, 7).

EXAMPLE

A factory manufactures two types of heavy-duty machines in quantities x and y . The joint-cost function is given by

$$f(x, y) = x^2 + 2y^2 - xy$$

To minimize cost, how many machines of each type should be produced if there must be a total of 8 machines?

The constraint is $x + y = 8$ and the objective function is

$$F(x, y, \lambda) = x^2 + 2y^2 - xy - \lambda(x + y - 8)$$

$$\frac{\partial F}{\partial x} = 2x - y - \lambda = 0$$

$$\frac{\partial F}{\partial y} = 4y - x - \lambda = 0$$

$$\frac{\partial F}{\partial \lambda} = -(x + y - 8)$$

Eliminating λ from the first two equations,

$$3x - 5y = 0$$

Solving simultaneously with the third equation,

$$3x + 3y = 24$$

$$8y = 24$$

$$y = 3$$

$$x = 5$$

$$\lambda = 7$$

$$\frac{\partial^2 F}{\partial x^2} = 2$$

$$\frac{\partial^2 F}{\partial y^2} = 4$$

$$\frac{\partial^2 F}{\partial x \partial y} = -1$$

$$\Delta = (2)(4) - (-1)^2 > 0$$

$$\frac{\partial^2 F}{\partial x^2} > 0 \quad \text{and} \quad \frac{\partial^2 F}{\partial y^2} > 0$$

so (5, 3) is a minimum.

NOTE: Assuming nonnegative cost, the unconstrained minimum of the function $f(x, y) = x^2 + 2y^2 - xy$ occurs at (0, 0).

Alternative method: $f(x, y) = x^2 + 2y^2 - xy$, $x + y = 8$.

$$\begin{aligned} G(x) &= x^2 + 2(8 - x)^2 - x(8 - x) \\ &= x^2 + 128 - 32x + 2x^2 - 8x + x^2 \\ &= 4x^2 - 40x + 128 \end{aligned}$$

$$\frac{dG}{dx} = 8x - 40 = 0 \quad \text{if } x = 5$$

$$y = 3$$

$$\frac{d^2 G}{dx^2} = 8 > 0$$

so (5, 3) is a minimum.

EXAMPLE

The relationship between sales S and the amounts x and y spent on two advertising media is given by

$$S = \frac{200x}{5 + x} + \frac{100y}{10 + y}$$

Net profit is $\frac{1}{5}$ of sales minus the cost of advertising. The advertising budget is 25; determine how it should be allocated between the two media in order to maximize net profit.

The constraint is $x + y = 25$ and the objective function is

$$\begin{aligned} F(x, y, \lambda) &= \frac{1}{5} \left[\frac{200x}{5 + x} + \frac{100y}{10 + y} \right] - x - y - \lambda(x + y - 25) \\ &= \frac{40x}{5 + x} + \frac{20y}{10 + y} - x - y - \lambda(x + y - 25) \\ \frac{\partial F}{\partial x} &= \frac{40(5 + x) - 40x}{(5 + x)^2} - 1 - \lambda = \frac{200}{(5 + x)^2} - 1 - \lambda = 0 \\ \frac{\partial F}{\partial y} &= \frac{20(10 + y) - 20y}{(10 + y)^2} - 1 - \lambda = \frac{200}{(10 + y)^2} - 1 - \lambda = 0 \\ \frac{\partial F}{\partial \lambda} &= -(x + y - 25) = 0 \end{aligned}$$

Eliminating λ from the first two equations,

$$\begin{aligned} \frac{200}{(5 + x)^2} &= \frac{200}{(10 + y)^2} \\ (5 + x)^2 &= (10 + y)^2 \end{aligned}$$

Solving simultaneously with the third equation,

$$x + y = 25$$

so

$$(5 + 25 - y)^2 = (10 + y)^2$$

$$30 - y = 10 + y$$

$$y = 10$$

$$x = 15$$

$$\lambda = -\frac{1}{2}$$

$$\frac{\partial^2 F}{\partial x^2} = 200(-2)(5 + x)^{-3} = -\frac{400}{(5 + x)^3}$$

$$\frac{\partial^2 F}{\partial y^2} = 200(-2)(10 + y)^{-3} = -\frac{400}{(10 + y)^3}$$

$$\frac{\partial^2 F}{\partial x \partial y} = 0$$

$$\Delta = \left(\frac{-400}{(5 + x)^3} \right) \left(\frac{-400}{(10 + y)^3} \right) - 0 > 0$$

$$\frac{\partial^2 F}{\partial x^2} < 0 \quad \text{and} \quad \frac{\partial^2 F}{\partial y^2} < 0$$

so (10, 15) is a maximum.

NOTE: The unconstrained maximum of the function

$$P(x, y) = \frac{1}{5} \left[\frac{200x}{5 + x} + \frac{100y}{10 + y} \right] - x - y \text{ occurs at } (10\sqrt{2} - 5, 10\sqrt{2} - 10).$$

$$\text{Alternative method: } P(x, y) = \frac{40x}{5 + x} + \frac{20y}{10 + y} - x - y, \quad x + y = 25.$$

$$P(x) = \frac{40x}{5 + x} + \frac{500 - 20x}{35 - x} - 25$$

$$\frac{dP}{dx} = \frac{40(5 + x) - 40x}{(5 + x)^2} + \frac{-20(35 - x) + (500 - 20x)}{(35 - x)^2}$$

$$= \frac{200}{(5 + x)^2} - \frac{200}{(35 - x)^2} = 0$$

$$\frac{200}{(5 + x)^2} = \frac{200}{(35 - x)^2}$$

$$5 + x = 35 - x$$

$$x = 15$$

$$y = 10$$

$$\frac{d^2 P}{dx^2} = 200(-2)(5 + x)^{-3} + 200(-2)(35 - x)^{-3}$$

$$= -\frac{400}{(5 + x)^3} - \frac{400}{(35 - x)^3} < 0$$

so (10, 15) is a maximum.

EXAMPLE

Maximum profit for a given production function and given prices of the inputs and output can be determined using Lagrange multipliers. Suppose the production function is

$$16z = 65 - 2(x - 5)^2 - 4(y - 4)^2$$

and the unit prices of the inputs x and y are 8 and 4, respectively, and the unit price of the output is 32. (See the example on page 363.) Profit is maximized if the net revenue function $32z - 8x - 4y$ is maximized subject to the production function constraint $16z - 65 + 2(x - 5)^2 + 4(y - 4)^2 = 0$.

$$F = 32z - 8x - 4y - \lambda[16z - 65 + 2(x - 5)^2 + 4(y - 4)^2]$$

$$\frac{\partial F}{\partial z} = 32 - 16\lambda = 0$$

$$\lambda = 2$$

$$\frac{\partial F}{\partial x} = -8 - 4(x - 5)\lambda = 0$$

$$x = 4$$

$$\frac{\partial F}{\partial y} = -4 - 8\lambda(y - 4) = 0$$

$$y = \frac{15}{4}$$

As shown previously, profit is maximized when $x = 4$, $y = \frac{15}{4}$. F is a function of three variables x , y , and z . The second-order conditions for maximization of functions of more than two variables, which are somewhat more complex than the conditions for functions of two variables, are discussed in Chapter 8.

PROBLEMS

Find the maxima and/or the minima of each of the following functions subject to the given constraint.

1. $f(x, y) = 3x^2 + 4y^2 - xy$ if $2x + y = 21$.
2. $f(x, y) = x^2 + y^2 - 2xy$ if $x^2 + y^2 = 50$.
3. $f(x, y) = x^2 - 10y^2$ if $x - y = 18$.
4. $f(x, y) = 3xy + 4y^2$ if $x^2 + y^2 = 10$.
5. $f(x, y) = x + y$ if $x^2 + y^2 = 1$.
6. $f(x, y) = x^2 + 24xy + 8y^2$ if $x^2 + y^2 = 25$.

Answers to Odd-Numbered Problems

1. minimum (8.5, 4)
3. maximum (20, 2)
5. minimum $(-\sqrt{\frac{1}{2}}, -\sqrt{\frac{1}{2}})$
maximum $(\sqrt{\frac{1}{2}}, \sqrt{\frac{1}{2}})$

KUHN-TUCKER CONDITIONS

The method of Lagrange multipliers can be modified to determine the maximum or minimum of a function of two variables subject to one inequality constraint as follows: Assume that the inequality constraint holds as an equality constraint and obtain the maximum or minimum using the method of Lagrange multipliers; if

$\lambda > 0$, this maximum or minimum is also the maximum or minimum subject to the inequality constraint; if $\lambda \leq 0$, the maximum or minimum determined without regard to the constraint satisfies the constraint and is thus also the constrained maximum or minimum.

This procedure can be generalized to include multiple inequality constraints, some of which may be satisfied as equalities and others as inequalities in any particular problem. In the more general procedure, the conditions necessary for a maximum or minimum subject to inequality constraints are known as the Kuhn-Tucker conditions. These conditions are given for the general case in Chapter 8; for the case of a function of two variables subject to one inequality constraint, the Kuhn-Tucker conditions are stated as follows.

A point (x, y) is a local maximum of $f(x, y)$ subject to $g(x, y) \leq 0$ only if there exists a nonnegative λ such that λ and (x, y) satisfy the following conditions:

$$\frac{\partial f(x, y)}{\partial x} - \lambda \frac{\partial g(x, y)}{\partial x} = 0$$

$$\frac{\partial f(x, y)}{\partial y} - \lambda \frac{\partial g(x, y)}{\partial y} = 0$$

$$\lambda g(x, y) = 0$$

$$g(x, y) \leq 0$$

These conditions are also sufficient if $f(x, y)$ is concave and the constraint is concave. Since a maximum point of $f(x, y)$ is a minimum point of $-f(x, y)$, this result is also applicable to minimizing a convex function subject to a convex constraint.

NOTE: A function $f(x, y)$ is convex in a region if a line segment drawn through any two points on the surface does not fall below the surface. A function of two variables is convex if and only if

$$f[(1-t)x_1 + tx_2, (1-t)y_1 + ty_2] \leq (1-t)f(x_1, y_1) + tf(x_2, y_2) \quad \text{for } 0 < t < 1$$

The function is strictly convex if \leq can be replaced by $<$; the function is concave if \leq can be replaced by \geq and strictly concave if \leq can be replaced by $>$.

A second-degree polynomial of the form

$$f(x, y) = Ax^2 + Bxy + Cy^2 + Dx + Ey + F$$

is convex if $4AC - B^2 > 0$ and $A > 0, C > 0$; it is concave if $4AC - B^2 > 0$ and $A < 0, C < 0$; and it is neither convex nor concave everywhere if $4AC - B^2 < 0$. Note that this result is a reformulation of the conditions for a maximum or minimum of two variables.

As shown in the following examples, the method of Lagrange multipliers modified for one inequality constraint and the Kuhn-Tucker conditions for one inequality constraint give the same solution. The advantage of the Kuhn-Tucker conditions is that they can be generalized for more than one inequality constraint. Although the method of Lagrange multipliers can be generalized for more than one equality constraint, it cannot conveniently be modified for more than one inequality constraint. If the function to be maximized or minimized is linear and the constraints are also linear, the solution may be obtained using linear programming for both equality and inequality constraints.

Note that the following examples are modifications of some of the examples for equality constraints.

EXAMPLE

Find the minimum of $f(x, y) = 5x^2 + 6y^2 - xy$ subject to the constraint $x + 2y \geq 24$.

Assuming the inequality constraint to hold as an equality constraint, $\lambda = 51$ (see page 370). Since $\lambda > 0$, the minimum, $x = 6$, $y = 9$, assuming the equality constraint is also the minimum assuming the inequality constraint.

Using the Kuhn-Tucker conditions,

$$\frac{\partial f}{\partial x} - \lambda \frac{\partial g}{\partial x} = 10x - y - \lambda = 0$$

$$\frac{\partial f}{\partial y} - \lambda \frac{\partial g}{\partial y} = 12y - x - 2\lambda = 0$$

$$\lambda g = \lambda(x + 2y - 24) = 0$$

$$g = x + 2y - 24 \geq 0$$

Either $\lambda = 0$ or $x + 2y - 24 = 0$. If $\lambda = 0$, then $x = y = 0$ in order to satisfy the first two equations, but then $x + 2y \geq 24$ is not satisfied. If $x + 2y - 24 = 0$, then $x = 24 - 2y$ and, substituting in the first two equations,

$$240 - 20y - y - \lambda = 0$$

$$12y - 24 + 2y - 2\lambda = 0$$

$$\lambda + 21y = 240$$

$$-2\lambda + 14y = 24$$

$$56y = 504$$

$$y = 9$$

$$x = 6$$

Thus, as above, the minimum of $f(x, y) = 5x^2 + 6y^2 - xy$ subject to $x + 2y \geq 24$ is $x = 6$, $y = 9$.

NOTE: $5x^2 + 6y^2 - xy$ is convex, since $4AC - B^2 = (4)(5)(6) - (-1)^2 > 0$ and $A > 0$, $C > 0$.

EXAMPLE

Find the maximum of $f(x, y) = 12xy - 3y^2 - x^2$ subject to the constraint $x + y \leq 16$.

Assuming the inequality constraint to hold as an equality constraint, $\lambda = 66$ (see page 371). Since $\lambda > 0$, the maximum, $x = 9$, $y = 7$, assuming the equality constraint is also the maximum subject to the inequality constraint.

Using the Kuhn-Tucker conditions,

$$\frac{\partial f}{\partial x} - \lambda \frac{\partial g}{\partial x} = 12y - 2x - \lambda = 0$$

$$\frac{\partial f}{\partial y} - \lambda \frac{\partial g}{\partial y} = 12x - 6y - \lambda = 0$$

$$\lambda g = \lambda(x + y - 16) = 0$$

$$g = x + y - 16 \leq 0$$

Either $\lambda = 0$ or $x + y - 16 = 0$. If $\lambda = 0$, then $x = y = 0$ in order to satisfy the first two equations, which satisfies $x + y \leq 16$; $f(0, 0) = 0$. If $x + y - 16 = 0$, then $x = 16 - y$ and, substituting in the first two equations,

$$12y - 32 + 2y - \lambda = 0$$

$$192 - 12y - 6y - \lambda = 0$$

$$14y - \lambda = 32$$

$$18y + \lambda = 192$$

$$32y = 224$$

$$y = 7$$

$$x = 9$$

The solution $x = 9$, $y = 7$ satisfies $12y - 2x - \lambda = 0$ and $12x - 6y - \lambda = 0$ if $\lambda = 66$. $f'(9, 7) = 528$, which is larger than $f(0, 0) = 0$. Thus, as above, the maximum of $f(x, y) = 12xy - 3y^2 - x^2$ subject to $x + y \leq 16$ is $x = 9$, $y = 7$.

NOTE: $12xy - 3y^2 - x^2$ is concave, since $4AC - B^2 = 4(-1)(-3) - (12)^2 < 0$ and $A < 0$, $C < 0$.

EXAMPLE

A factory manufactures two types of heavy-duty machines, x and y . The joint-cost function is given by

$$f(x, y) = x^2 + 2y^2 - xy$$

To minimize cost, how many machines of each type should be produced if there must be a total of at least 8 machines?

Assuming the inequality constraint to hold as an equality constraint, $\lambda = 7$ (see page 372). Since $\lambda > 0$, the maximum, $x = 5$, $y = 3$, assuming the equality constraint is also the maximum subject to the inequality constraint.

Using the Kuhn-Tucker conditions,

$$\frac{\partial f}{\partial x} - \lambda \frac{\partial g}{\partial x} = 2x - y - \lambda = 0$$

$$\frac{\partial f}{\partial y} - \lambda \frac{\partial g}{\partial y} = 4y - x - \lambda = 0$$

$$\lambda g = \lambda(x + y - 8) = 0$$

$$g = x + y - 8 \geq 0$$

Either $\lambda = 0$ or $x + y - 8 = 0$. If $\lambda = 0$, then $x = y = 0$ in order to satisfy the first two equations, but then $x + y \geq 8$ is not satisfied. If $x + y - 8 = 0$, then $x = 8 - y$ and, substituting in the first two equations,

$$16 - 2y - y - \lambda = 0$$

$$4y - 8 + \lambda = 0$$

$$3y + \lambda = 16$$

$$5y - \lambda = 8$$

$$8y = 24$$

$$y = 3$$

Thus, as above, the minimum of $f(x, y) = x^2 + 2y^2 - xy$ subject to $x + y \geq 8$ is $x = 5, y = 3$.

NOTE: $x^2 + 2y^2 - xy$ is convex, since $4AC - B^2 = 4(1)(2) - (-1)^2 > 0$ and $A > 0, C > 0$.

EXAMPLE

The relationship between sales S and the amounts x and y spent on two advertising media is given by

$$S = \frac{200x}{5+x} + \frac{100y}{10+y}$$

Net profit is $\frac{1}{5}$ of sales minus the cost of advertising. The advertising budget has a maximum of 20; determine how it should be allocated between the two media to maximize net profit.

Assuming the inequality constraint to hold as an equality constraint, $\lambda = -\frac{1}{2}$ (see page 373). Since $\lambda < 0$, the unconstrained maximum is also the maximum subject to the inequality constraint.

$$f(x, y) = \frac{1}{5} \left[\frac{200x}{5+x} + \frac{100y}{10+y} \right] - x - y$$

$$\frac{\partial f}{\partial x} = \frac{40(5+x) - 40x}{(5+x)^2} - 1$$

$$\frac{\partial f}{\partial y} = \frac{20(10+y) - 20y}{(10+y)^2} - 1$$

$$(5+x)^2 = 200$$

$$x = 10\sqrt{2} - 5$$

$$(10+y)^2 = 200$$

$$y = 10\sqrt{2} - 10$$

Using the Kuhn-Tucker conditions,

$$\frac{\partial f}{\partial x} - \lambda \frac{\partial g}{\partial x} = \frac{40(5+x) - 40x}{(5+x)^2} - 1 - \lambda = 0$$

$$\frac{\partial f}{\partial y} - \lambda \frac{\partial g}{\partial y} = \frac{20(10+y) - 20y}{(10+y)^2} - 1 - \lambda = 0$$

$$\lambda g = \lambda(x + y - 25) = 0$$

$$g(x, y) = x + y - 25 \leq 0$$

$$\text{NOTE: } f(x, y) = \frac{1}{5} \left[\frac{200x}{5+x} + \frac{100y}{10+y} \right] - x - y$$

Either $\lambda = 0$ or $x + y - 25 = 0$. If $\lambda = 0$, then in order to satisfy the first two equations,

$$40(5+x) - 40x = (5+x)^2$$

$$(5+x)^2 = 200$$

and

$$x = 10\sqrt{2} - 5$$

$$20(10+y) - 20y = (10+y)^2$$

$$(10+y)^2 = 200$$

$$y = 10\sqrt{2} - 10$$

$x = 10\sqrt{2} - 5$, $y = 10\sqrt{2} - 10$, satisfies $x + y \leq 25$. $f(10\sqrt{2} - 5, 10\sqrt{2} - 10) = 75 - 40\sqrt{2} \approx 18.44$. If $x + y - 25 = 0$, $x = 25 - y$, and (see page 374)

$$(5 + 25 - y)^2 = (10 + y)^2$$

$$30 - y = 10 + y$$

$$y = 10$$

$$x = 15$$

The solution $x = 15$, $y = 10$, satisfies

$$\frac{200}{(5 + x)^2} - 1 - \lambda = 0$$

$$\frac{200}{(10 + y)^2} - 1 - \lambda = 0$$

if $\lambda = -\frac{1}{2}$. $f(15, 10) = 15$, which is less than $f(10\sqrt{2} - 5, 10\sqrt{2} - 10) = 75 - 40\sqrt{2}$. Thus, as above, the maximum of

$$f(x, y) = \frac{1}{5} \left[\frac{200x}{5 + x} + \frac{100y}{10 + y} \right] - x - y$$

subject to $x + y \leq 25$ is $x = 10\sqrt{2} - 5$, $y = 10\sqrt{2} - 10$.

3.6 Applications of Constrained Maximization and Minimization in Business and Economics

There are many applications of constrained maxima and minima in business and economics, as illustrated in the examples of the preceding section. One of the most important applications of constrained maximization in economics occurs in utility theory; the consumer maximizes the utility derived from the purchase of two commodities, where the consumer's fixed income and the prices of the commodities constitute a budget constraint. The problem of constrained utility maximization is considered in this section.

CONSTRAINED UTILITY MAXIMIZATION

If his purchases are limited to two commodities Q_1 and Q_2 , the rational consumer wishes to purchase that combination of Q_1 and Q_2 which will maximize his utility, subject to his budget constraint. Since his purchases are limited to the two commodities, the consumer's budget constraint can be written

$$y^\circ = p_1 q_1 + p_2 q_2$$

where y° is his fixed income and p_1 , p_2 , q_1 , and q_2 are the prices and quantities of Q_1 and Q_2 , respectively.

The budget line $y^\circ = p_1 q_1 + p_2 q_2$ shows all possible combinations of q_1 and q_2 which the consumer can purchase, given his budget constraint. In order to maximize his utility, the consumer will choose to purchase a combination of q_1 and q_2 lying on the highest possible indifference curve. As shown in Figure 3.5, this maximizing point

is the point at which the budget line is tangent to an indifference curve. Any movement from this point along the budget constraint puts the consumer on a lower indifference curve and results in diminished utility.

Mathematically, the problem is one of obtaining a constrained maximum:

$$\text{maximize } U(q_1, q_2)$$

$$\text{subject to } y^\circ = p_1 q_1 + p_2 q_2$$

The budget constraint can be written

$$\frac{y^\circ - p_1 q_1}{p_2} = q_2$$

and the utility function, as a function only of q_1 , is

$$U = f\left(q_1, \frac{y^\circ - p_1 q_1}{p_2}\right)$$

Note that q_2 is determined if q_1 is specified, so it is sufficient to maximize U with respect to q_1

$$\frac{dU}{dq_1} = f_1 + f_2 \left(-\frac{p_1}{p_2}\right) = 0$$

and

$$\frac{f_1}{f_2} = \frac{p_1}{p_2}$$

NOTE: The utility function may be written

$$U = f(q_1, q_2) \quad \text{where } q_2 = \frac{y^\circ - p_1 q_1}{p_2} = g(q_1)$$

Then

$$\begin{aligned} \frac{dU}{dq_1} &= \frac{\partial f}{\partial q_1} \cdot \frac{\partial q_1}{\partial q_1} + \frac{\partial f}{\partial q_2} \cdot \frac{\partial q_2}{\partial q_1} \\ &= f_1 + f_2 \left(-\frac{p_1}{p_2}\right) \end{aligned}$$

Thus, for a maximum, the ratio of the marginal utilities must equal the ratio of the prices, or equivalently,

$$\frac{f_1}{p_1} = \frac{f_2}{p_2}$$

that is, marginal utility divided by price must be the same for both commodities. This condition is necessary, but not sufficient, for occurrence of a maximum. The second-order condition for a maximum requires that

$$\frac{d^2 U}{dq_1^2} < 0$$

But

$$\begin{aligned}\frac{\partial^2 U}{\partial q_1^2} &= \frac{\partial}{\partial q_1} \left(\frac{dU}{dq_1} \right) \left(\frac{\partial q_1}{\partial q_1} \right) + \frac{\partial}{\partial q_2} \left(\frac{dU}{dq_1} \right) \left(\frac{\partial q_2}{\partial q_1} \right) \\ &= \left[f_{11} + f_{12} \left(-\frac{p_1}{p_2} \right) \right] + f_{12} \left(\frac{p_1}{p_2} \right) + f_{22} \left(-\frac{p_1}{p_2} \right)^2 \\ &= f_{11} + 2f_{12} \left(-\frac{p_1}{p_2} \right) + f_{22} \left(-\frac{p_1}{p_2} \right)^2\end{aligned}$$

The second-order condition for a maximum requires that

$$f_{11} + 2f_{12} \left(-\frac{p_1}{p_2} \right) + f_{22} \left(-\frac{p_1}{p_2} \right)^2 < 0$$

or

$$f_{11} p_2^2 - 2f_{12} p_1 p_2 + f_{22} p_1^2 < 0$$

In Technical Note I, it is shown that the second-order condition for a maximum requires that the utility curves are convex from below. Since $\frac{dq_2}{dq_1} = -\frac{f_1}{f_2} = -\frac{p_1}{p_2}$ and prices are positive, the indifference curves are negatively sloped. Thus if the utility function has a maximum, the indifference curves are of the general form shown in Figure 3.5.

EXAMPLE

If $U = q_1 q_2$, $p_1 = 15$, $p_2 = 5$, and the consumer's income for the period is 150, the budget constraint is

$$150 - 15q_1 - 5q_2 = 0$$

$$q_2 = \frac{150 - 15q_1}{5} = 30 - 3q_1$$

and

$$U = 30q_1 - 3q_1^2$$

$$\frac{dU}{dq_1} = 30 - 6q_1$$

If $\frac{dU}{dq_1} = 0$, then $q_1 = 5$, $q_2 = 15$, and $\frac{d^2 U}{dq_1^2} = -6$, so the consumer maximizes utility subject to his budget constraint by purchasing 5 units of Q_1 and 15 units of Q_2 . Note that

$$\frac{f_1}{f_2} = \frac{q_2}{q_1} = \frac{15}{5} = \frac{p_1}{p_2}$$

as required for a maximum.

EXAMPLE

If $U = q_1 q_2^2 - 10q_1$, $p_1 = 2$, $p_2 = 8$, and the consumer's income for the period is 116, the budget constraint is

$$116 - 2q_1 - 8q_2 = 0$$

$$q_1 = \frac{116 - 8q_2}{2} = 58 - 4q_2$$

and

$$\begin{aligned} U &= (58 - 4q_2)q_2^2 - 10(58 - 4q_2) \\ &= 58q_2^2 - 4q_2^3 - 580 + 40q_2 \end{aligned}$$

$$\frac{dU}{dq_2} = 116q_2 - 12q_2^2 + 40$$

If $\frac{dU}{dq_2} = 0$,

$$3q_2^2 - 29q_2 - 10 = 0$$

$$(3q_2 + 1)(q_2 - 10) = 0$$

$$q_2 = -\frac{1}{3}, 10$$

But $q_2 \geq 0$, so the solution is $q_2 = 10$, $q_1 = 18$.

$$\frac{d^2U}{dq_2^2} = 116 - 24q_2$$

$$\left. \frac{d^2U}{dq_2^2} \right|_{q_2=10} = 116 - 240 = -124 < 0$$

So the consumer maximizes utility subject to his budget constraint by purchasing 18 units of Q_1 and 10 units of Q_2 .

Note that

$$\frac{f_1}{f_2} = \frac{q_2^2 - 10}{2q_1 q_2} = \frac{90}{360} = \frac{2}{8} = \frac{p_1}{p_2}$$

as required for a maximum.

Alternatively, the constrained maximum can be obtained using Lagrange multipliers

$$Z = f(q_1, q_2) - \lambda(y^\circ - p_1 q_1 - p_2 q_2)$$

$$\frac{\partial Z}{\partial q_1} = f_1 + \lambda p_1 = 0$$

$$\frac{\partial Z}{\partial q_2} = f_2 + \lambda p_2 = 0$$

$$\frac{\partial Z}{\partial \lambda} = y^\circ - p_1 q_1 - p_2 q_2 = 0$$

Thus

$$\lambda = -\frac{f_1}{p_1} = -\frac{f_2}{p_2}$$

and

$$\frac{f_1}{f_2} = \frac{p_1}{p_2}$$

and so forth, as above.

PROBLEMS

1. Find the minimum of $f(x, y) = 4x^2 + 5y^2 - 6y$ if $x + 2y \geq 18$.
2. Find the maximum of $f(x, y) = 16x + 12y - 2x^2 - 3y^2$ if $x + y \leq 11$.
3. Find the minimum of $f(x, y) = 3x^2 + 3y^2$ if $x + y \geq 10$.
4. Find the minimum of $f(x, y) = 12x^2 + 4y^2 - 8xy - 32x$ if $x + y \geq 1$.
5. Find the maximum of $f(x, y) = 10xy - 5x^2 - 7y^2 + 40x$ if $x + y \leq 13$.
6. Find the maximum of $f(x, y) = 6xy - 3x^2 - 4y^2$ if $3x + y \leq 19$.
7. Production, P , as a function of two inputs x and y , is given by

$$P = x^2 + 5xy - 4y^2$$

Find the amounts of x and y which maximize production if a. $2x + 3y = 74$, b. $2x + 3y \leq 74$.

8. Sales, S , as a function of the amounts, x and y , spent on two types of promotion is given by

$$S = \frac{240x}{25 + 3x} + \frac{150y}{10 + y}$$

Net profit $= \frac{1}{10}S - x - y$. Find the allocation of x and y which will maximize net profit if $x + y = 15$.

9. Cost of production, C , as a function of the numbers produced, x and y , of two types of items is given by

$$C = 6x^2 + 3y^2$$

To minimize cost, what numbers of the two items should be produced if a. $x + y = 18$, b. $x + y \geq 18$?

10. The cost of repairs, C , as a function of the numbers, x and y , of inspections at two points in a process is given by

$$C = 2x^2 + 3y^2 + xy - 22x + 5$$

To minimize repair cost, what number of inspections should be made at each point if $x - y = 2$?

11. The number of breakdowns, N , as a function of the numbers, x and y , of replacements of two parts of a machine is given by

$$N = 3x^2 + y^2 + 2xy - 22x + 6$$

To minimize breakdowns, what numbers of replacements should be made for each part if $2x = y$?

12. The cost of repairs, C , as a function of the numbers, x and y , of inspections at two points in a process is given by

$$C = 4x^2 + 2y^2 + 5xy - 20x + 30$$

To minimize repair costs, what number of inspections should be made at each point if the total number of inspections is a. 10, b. not fewer than 10?

13. Using Lagrange multipliers, determine the maximum profit if the production function is $z = 20 - x^2 + 10x - 2y^2 + 5y$, the prices of the inputs x and y are 2 and 1, respectively, and the price of the output is 5. (See Problem 21, page 366.)
14. Using Lagrange multipliers, determine the maximum profit if the production function is $z = 10 - 2x^2 + xy - y^2 + 5y$, the prices of the inputs x and y are each 3, and the price of the output is 6. (See Problem 22, page 366.)
15. If the consumer's utility function is $U = q_1^2 q_2$, $p_1 = 4$, $p_2 = 5$, and $y^\circ = 120$, determine the quantities q_1 and q_2 which he should purchase in order to maximize his derived utility.
16. If the consumer's utility function is $U = q_1 q_2 - q_1^2$, $p_1 = 3$, $p_2 = 6$, and $y^\circ = 90$, determine the quantities q_1 and q_2 which he should purchase in order to maximize his derived utility.
17. If the consumer's utility function is $U = q_1^2 + 2q_2^2 + 5q_1 q_2$, $p_1 = 5$, $p_2 = 10$, and the consumer's income for the period is 90, determine the quantities q_1 and q_2 which he should purchase in order to maximize his derived utility.

18. If the consumer's utility function is $U = q_1 q_2 - 3q_2^2$, $p_1 = 10$, $p_2 = 15$, and the consumer's income for the period is 180, determine the quantities q_1 and q_2 which he should purchase in order to maximize his derived utility.

Answers to Odd-Numbered Problems

- | | |
|-----------------------|--|
| 1. (4, 7) | 11. $x = 1, y = 2$ |
| 3. (5, 5) | 13. $P_{\max} = 229\frac{3}{5}, x = \frac{24}{5}, y = \frac{6}{5}$ |
| 5. (8, 5) | 15. $q_1 = 20, q_2 = 8$ |
| 7. a. $x = 31, y = 4$ | 17. $q_1 = \frac{27}{2}, q_2 = \frac{9}{4}$ |
| b. $x = 31, y = 4$ | |
| 9. a. $x = 6, y = 12$ | |
| b. $x = 6, y = 12$ | |

Technical Note I. Proof of Convexity of Utility Curves

Obtain the total derivative $\frac{d^2 q_2}{dq_1^2}$ by further differentiating $\frac{dq_2}{dq_1} = -\frac{f_1}{f_2}$

$$\frac{d^2 q_2}{dq_1^2} = \left[\frac{\partial}{\partial q_1} \left(\frac{dq_2}{dq_1} \right) \right] \left[\frac{\partial q_1}{\partial q_1} \right] + \left[\frac{\partial}{\partial q_2} \left(\frac{dq_2}{dq_1} \right) \right] \left[\frac{\partial q_2}{\partial q_1} \right]$$

But

$$\frac{\partial}{\partial q_1} \left(\frac{dq_2}{dq_1} \right) = \frac{\partial}{\partial q_1} \left(-\frac{f_1}{f_2} \right) = -\frac{f_{11} f_2 - f_1 f_{12}}{f_2^2}$$

$$\frac{\partial q_1}{\partial q_1} = 1$$

$$\frac{\partial}{\partial q_2} \left(\frac{dq_2}{dq_1} \right) = \frac{\partial}{\partial q_2} \left(-\frac{f_1}{f_2} \right) = -\frac{f_{12} f_2 - f_1 f_{22}}{f_2^2}$$

$$\frac{\partial q_2}{\partial q_1} = -\frac{f_1}{f_2}$$

Thus

$$\begin{aligned} \frac{d^2 q_2}{dq_1^2} &= -\frac{f_{11} f_2 - f_1 f_{12} + (f_{12} f_2 - f_1 f_{22}) \left(-\frac{f_1}{f_2} \right)}{f_2^2} \\ &= -\frac{1}{f_2^3} (f_{11} f_2^2 - 2f_{12} f_1 f_2 + f_{22} f_1^2) \end{aligned}$$

Substituting $f_1 = (p_1/p_2)/f_2$,

$$\begin{aligned} \frac{d^2 q_2}{dq_1^2} &= -\frac{1}{f_2^3} \left(f_{11} f_2^2 - 2f_{12} f_2^2 \frac{p_1}{p_2} + f_{22} f_2^2 \frac{p_1^2}{p_2^2} \right) \\ &= -\frac{1}{f_2 p_2^2} (f_{11} p_2^2 - 2f_{12} p_1 p_2 + f_{22} p_1^2) \end{aligned}$$

But the second-order condition for a maximum requires that

$$f_{11} p_2^2 - 2f_{12} p_1 p_2 + f_{22} p_1^2 < 0$$

and, since $f_2 > 0$,

$$\frac{d^2 q_2}{dq_1^2} > 0$$

and the indifference curves are convex from below.

Integral Calculus

4

4.1 Introduction

Integration has two distinct interpretations; integration is a procedure which is the inverse of differentiation, and it is a method of determining the area under a curve. There are many important applications of integration in each of these contexts.

As an operation, integration is the inverse of differentiation. Thus if a function is differentiated and the resulting function is then integrated, the result is the original function. As discussed below, this is precisely true only if the constant of integration is specified in some way; otherwise the result may differ from the original function by a constant. In this context, integration is the process of finding a function when its derivative (or rate of change) is known. In economics, integration can be used to find a total cost function when the marginal cost function is given, to find a total revenue function when the marginal revenue function is given, and so forth.

Integration can also be defined as the process of finding the limiting value of a sum of terms when the number of terms increases infinitely and the numerical value of each term approaches zero. It is in this context that integration is interpreted as finding the area under a curve. In fact, integral calculus was developed for the purpose of evaluating areas by supposing them to be divided into an infinite number of infinitesimally small parts whose sum is the area required; the integral sign is the elongated S used by early writers to indicate "sum." In economics total revenue can be evaluated as the area under the marginal revenue curve, consumer's surplus and producer's surplus can be evaluated as areas under demand and supply curves, and so forth.

For either application, integration requires operationally that a function be determined when its derivative is given. Unfortunately, the techniques of integration are inherently more difficult than those of differentiation and there are functions, some of them deceptively simple in appearance, whose integrals cannot be expressed in terms of elementary functions. The simpler cases of integration are accomplished by reversing the corresponding formulas for differentiation; more complicated cases are handled by the use of tables of standard forms, by various procedures of substitution, and, if necessary, by numerical (approximation) methods.

The two interpretations of integration and some of their applications in business and economics are discussed first, using functions that are easily integrated by revers-

ing the formulas for differentiation. Subsequent sections discuss special methods useful for integrating more difficult functions and several approximate methods of integration. Multiple integration, the inverse of partial differentiation, is also defined and illustrated.

4.2 Indefinite Integration

The process of determining a function whose derivative is known is called *integration* and the required function is called an *integral* or an *antiderivative* of the given function. Integration of a function is defined and illustrated in this section.

If $F(x)$ is an integral with respect to x of the function $f(x)$, the relationship between $F(x)$ and $f(x)$ is expressed as follows:

$$\int f(x) dx = F(x) + C$$

where the left-hand member is read “integral of f of x with respect to x .” The symbol \int is an *integral sign*, $f(x)$ is the *integrand*, $F(x)$ is a *particular integral*, C is the *constant of integration*, and $F(x) + C$ is the *indefinite integral*.

Note that if $F(x)$ is an integral of $f(x)$ with respect to x , then $F(x) + C$ is also such an integral, where C is any constant, since the derivative of any constant is zero—that is,

$$\begin{aligned} \frac{d}{dx}[F(x) + C] &= \frac{dF(x)}{dx} + \frac{dC}{dx} \\ &= \frac{dF(x)}{dx} \end{aligned}$$

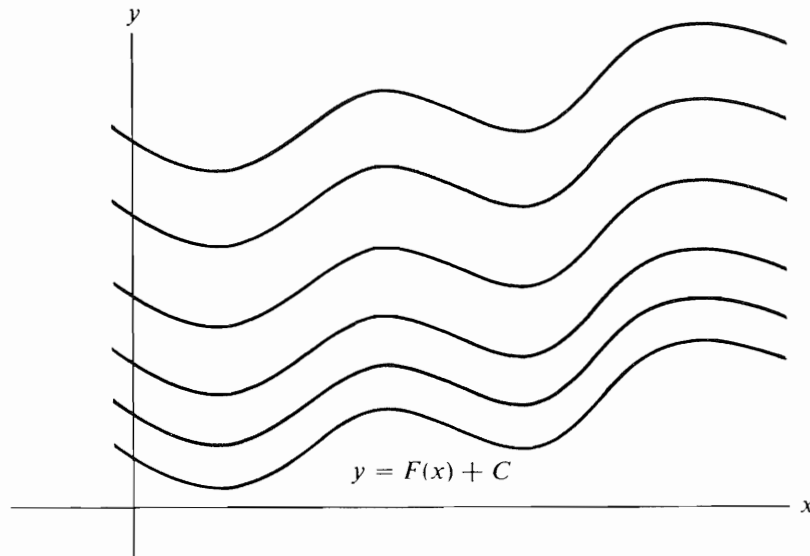
which is $f(x)$ if $F(x)$ is an integral of $f(x)$ with respect to x . Thus a function whose derivative is given is not completely determined, since the integral includes an arbitrary additive constant, the constant of integration. It is for this reason that the function $\int f(x) dx$ is referred to as the *indefinite integral* of $f(x)$.

It can be shown that two functions which have the same derivative differ at most by a constant; thus if $F(x)$ is an integral of $f(x)$, all integrals of $f(x)$ are included in the set $F(x) + C$, where C is any constant. In many applications of integration, certain information given in the problem, often referred to as an *initial* or *boundary condition*, uniquely specifies the constant of integration.

Geometrically, $y = F(x) + C$ represents a family of curves any one of which can be obtained by shifting the curve $y = F(x)$ (corresponding to $C = 0$) through a vertical displacement C (see Figure 4.1). The curves represented by $y = F(x) + C$ are parallel to each other in the sense that the slope of the tangent to any one of them at the point having abscissa x is $f(x)$. Thus this family of curves has the property that, given any point (x_0, y_0) , there is one and only one curve of the family which passes through this particular point. For the curve to pass through the point, its equation must be satisfied by the coordinates of the point, and this uniquely specifies the value of C , since then

$$C = y_0 - F(x_0)$$

Figure 4.1



With C thus determined, a definite function expressing y in terms of x is obtained, that is, the constant of integration is uniquely determined if one point is specified through which the curve representing the integral passes. This specification is referred to as an *initial condition*, because evaluation of the constant of integration was first done in connection with problems in mechanics where initial velocities or positions of moving bodies are specified; however, the terminology is usually appropriate for business or behavioral-science applications as well, since the point most likely to be specified in such cases is the origin or an intercept.

EXAMPLE

The integral of $f(x) = 4x - 3$ can be obtained by reversing the appropriate formulas for differentiation:

$$\int (4x - 3) dx = 2x^2 - 3x + C$$

Note that $\frac{d}{dx}(2x^2 - 3x + C) = 4x - 3$, regardless of the value of C . Thus, geometrically, the integral is represented by a family of parabolas each corresponding to a different value of C . (See Figure 4.2.) Suppose it is specified, for example, that $f(x) = 5$ when $x = 0$; then

$$2x^2 - 3x + C = 5$$

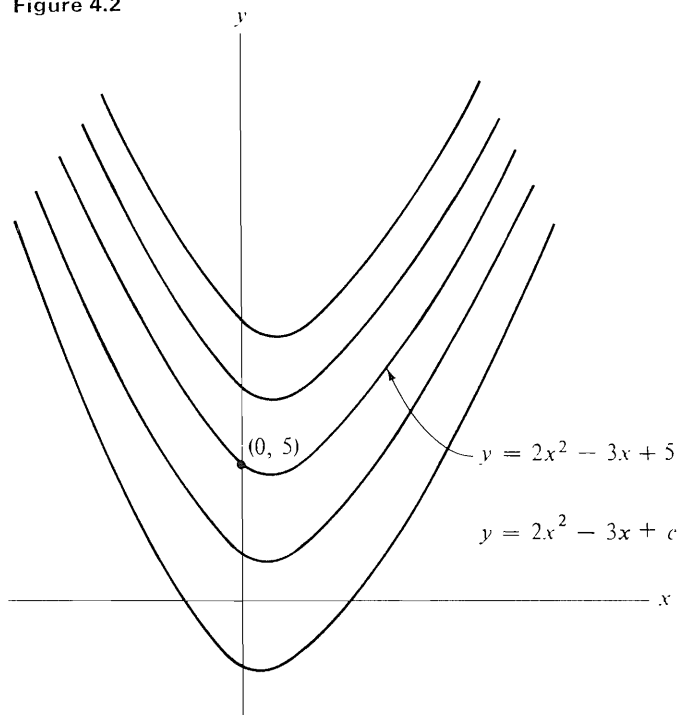
$$C = 5$$

and

$$F(x) = 2x^2 - 3x + 5$$

which specifies a particular curve of the family given by $F(x) = 2x^2 - 3x + C$. (See Figure 4.2.)

Figure 4.2



The rules (or standard forms) for integration used in the preceding example are obtained directly by reversing the corresponding rules for differentiation and can be summarized as follows:

1. $\int dx = x + C$.
2. $\int K dx = K \int dx$, where K is any constant.
3. $\int (du + dv) = \int du + \int dv$, where $u = f(x)$ and $v = g(x)$ are differentiable functions of x .
4. $\int x^n dx = \frac{x^{n+1}}{n+1} + C$, $n \neq -1$.
5. $\int u^n du = \frac{u^{n+1}}{n+1} + C$, $n \neq -1$,

where $u = f(x)$ is a differentiable function of x .

EXAMPLE

$$\int (8x^2 + 4) dx = \frac{8x^3}{3} + 4x + C \quad (\text{rules 1, 2, 3, 4})$$

$$\int x^{3/2} dx = \frac{2x^{5/2}}{5} + C \quad (\text{rule 4})$$

$$\int (3x^2 + 2)^3 x \, dx = \frac{1}{6} \int \underbrace{(3x^2 + 2)^3}_{u^3} \underbrace{(6x) \, dx}_{du} \quad (\text{rule 2})$$

$$= \left(\frac{1}{6}\right) \frac{(3x^2 + 2)^4}{4} + C \quad (\text{rule 5})$$

$$= \frac{(3x^2 + 2)^4}{24} + C$$

Alternatively,

$$\begin{aligned} \int (3x^2 + 2)^3 x \, dx &= \int (27x^7 + 54x^5 + 36x^3 + 8x) \, dx \\ &= \frac{27x^8}{8} + 9x^6 + 9x^4 + 4x^2 + C \\ &= \frac{81x^8 + 216x^6 + 216x^4 + 96x^2}{24} + C \\ &= \frac{(3x^2 + 2)^4}{24} + C \end{aligned}$$

Note that C in the last step is not equal to C in the preceding step, since

$$(3x^2 + 2)^4 = 81x^8 + 216x^6 + 216x^4 + 96x^2 + 16$$

and the value of C in the last step is thus $\frac{16}{24}$ less than the value of C in the preceding step. It is customary to make no notational distinction when the value of a constant of integration is changed by changing the form of the equation.

EXAMPLE

Evaluate the following integrals.

$$\begin{aligned} \text{(a)} \quad y &= \int (x^3 + 2x^{5/2} + 5x^{3/2} + 10x) \, dx \quad \text{if } y = 0 \text{ when } x = 0 \\ &= \frac{1}{4}x^4 + \frac{4}{7}x^{7/2} + 2x^{5/2} + 5x^2 + C \end{aligned}$$

If $x = 0$, $y = 0$, then $C = 0$ and

$$y = \frac{1}{4}x^4 + \frac{4}{7}x^{7/2} + 2x^{5/2} + 5x^2$$

$$\text{(b)} \quad y = \int (x + 2)^2 \, dx \quad \text{if } y = 10 \text{ when } x = 1$$

$$= \frac{(x + 2)^3}{3} + C$$

If $x = 1$, $y = 10$, then $C = 1$ and

$$y = \frac{(x + 2)^3}{3} + 1$$

$$\text{(c)} \quad y = \int (3x^2 + 2x + 6) \, dx \quad \text{if } y = 5 \text{ when } x = 0$$

$$= x^3 + x^2 + 6x + C$$

If $x = 0$, $y = 5$, then $C = 5$ and

$$y = x^3 + x^2 + 6x + 5$$

EXAMPLE

For each of the following, find the curve having the given slope and passing through the given point.

$$(a) \frac{dy}{dx} = \frac{x^2 - 4}{x^2}, \quad (4, 1)$$

$$(b) \frac{dy}{dx} = x^2\sqrt{x}, \quad (1, 0)$$

$$\begin{aligned} (a) \int \frac{x^2 - 4}{x^2} dx &= \int dx - \int \frac{4}{x^2} dx \\ &= x + \frac{4x^{-1}}{1} + C \\ &= x + \frac{4}{x} + C \end{aligned}$$

Thus

$$y = x + \frac{4}{x} + C$$

$$1 = 4 + 1 + C$$

$$C = -4$$

so

$$\begin{aligned} y &= x + \frac{4}{x} - 4 \\ &= \frac{x^2 - 4x + 4}{x} \\ &= \frac{(x - 2)^2}{x} \end{aligned}$$

is the required curve.

$$\begin{aligned} (b) \int x^2\sqrt{x} dx &= \int x^{5/2} dx \\ &= \frac{x^{7/2}}{\frac{7}{2}} + C \\ &= \frac{2}{7}x^{7/2} + C \end{aligned}$$

Thus

$$y = \frac{2}{7}x^{7/2} + C$$

$$0 = \frac{2}{7} + C$$

$$C = -\frac{2}{7}$$

so

$$\begin{aligned} y &= \frac{2}{7}x^{7/2} - \frac{2}{7} \\ &= \frac{2}{7}(x^{7/2} - 1) \\ &= \frac{2}{7}(x^3\sqrt{x} - 1) \end{aligned}$$

is the required curve.

PROBLEMS

Evaluate the following integrals.

1. $\int (x^2 - \sqrt{x} + 4) dx$

9. $\int \frac{(1 - 2x)^2 dx}{\sqrt{2x}}$

2. $\int (2 - 7t)^{2/3} dt$

10. $\int \frac{dx}{x\sqrt{2x}}$

3. $\int \sqrt{2 + 5y} dy$

11. $\int (x\sqrt{x} - 5)^2 dx$

4. $\int \frac{dx}{(3x + 2)^2}$

12. $\int \frac{(x^3 - 1) dx}{x - 1}$

5. $\int \frac{3r dr}{\sqrt{1 - r^2}}$

13. $\int (2x + 3) dx$

6. $\int x\sqrt{2x^2 + 1} dx$

14. $\int (x^2 - \sqrt{x}) dx$

7. $\int \left(\sqrt{x} + \frac{1}{\sqrt{x}} \right) dx$

15. $\int (\sqrt{2 + 5y}) dy$

8. $\int \frac{(z + 1) dz}{\sqrt[3]{z^2 + 2z + 2}}$

16. Find the equation of the curve having slope $\frac{dy}{dx} = 2x - 5$ and passing through the point (5, 4).

17. Find the equation of the curve having slope $\frac{dy}{dx} = (x + 1)(x + 2)$ and passing through the point $(-3, -\frac{3}{2})$.

18. If $\frac{dy}{dx} = 2x - 3$ and $y = 2$ if $x = 3$, find the value of y if $x = 5$.

19. If $\frac{dp}{dx} = \frac{1}{\sqrt{2ax}}$ and $p = 2a$ if $x = \frac{1}{2}a^3$, find the value of p if $x = 2a^3$.

20. Find the equation of the curve for which $y''' = 2$ and which has a slope of -2 at its point of inflection (1, 3).

21. Find the equation of the curve for which $y'' = x^3$ and which is tangent to the line $2x + y = 5$ at the point (1, 3).

22. Find the equation of the curve for which $y'' = 6x^2$ and which passes through the points (0, 2) and (-1, 3).

23. Find the equation of the curve for which $y'' = x$ and which passes through the point (1, 2) with a slope of $\frac{5}{2}$.

24. Find the equation of the curve which has a slope of zero at the point (0, 2), has a point of inflection at $(-1, \frac{10}{3})$, and for which $y''' = 4$.

Answers to Odd-Numbered Problems

1. $\frac{1}{3}x^3 - \frac{2}{5}x^{5/2} + 4x + C$

9. $\sqrt{2x} (1 - \frac{4}{3}x + \frac{4}{5}x^2) + C$

3. $\frac{2}{15}(2 + 5y)^{3/2} + C$

11. $\frac{1}{4}x^4 - 4x^{5/2} + 25x + C$

5. $-3\sqrt{1 - r^2} + C$

13. $x^2 + 3x + C$

7. $\frac{2}{3}x^{3/2} + 2x^{1/2} + C$

15. $\frac{2}{15}(2 + 5y)^{3/2} + C$

17. $y = \frac{1}{3}x^3 + \frac{3}{2}x^2 + 2x$

19. $p(2a^3) = 3a$

21. $y = \frac{2}{x} + 1$

23. $y = \frac{1}{6}x^3 + 2x - \frac{1}{6}$

4.3 Applications of Indefinite Integration in Business and Economics

As mentioned in previous discussions of differentiation, in economics the variation of one quantity y with respect to another quantity x is frequently discussed in terms of two concepts—*average* variation and *marginal* variation. Just as its marginal variation can be obtained by differentiating a function, the function (apart from a constant) can be obtained by integrating its marginal variation. This application of integration is illustrated for cost, revenue, and consumption functions in this section. A similar application is illustrated for capital formation.

COST

If the total cost y of producing and marketing x units of a commodity is given by the function

$$y = f(x)$$

Then the *average cost* per unit is

$$\frac{y}{x} = \frac{f(x)}{x}$$

and the *marginal cost* is

$$y' = \frac{dy}{dx} = f'(x)$$

That is, marginal cost is the derivative with respect to x , $f'(x)$, of the total cost function $y = f(x)$. Thus total cost is the integral with respect to x of the marginal cost function; that is,

$$y = \int f'(x) dx = f(x) + C$$

To obtain a unique total cost function by integrating the corresponding marginal cost function, an initial condition must be specified. Frequently this specification is in terms of a fixed cost or initial overhead, that is, the cost when $x = 0$.

EXAMPLE

Marginal cost y' as a function of units produced x is given by

$$y' = 1.064 - 0.005x$$

Find the total and average cost functions if fixed cost is 16.3.

$$\begin{aligned} y &= \int (1.064 - 0.005x) dx \\ &= 1.064x - 0.0025x^2 + C \end{aligned}$$

If $x = 0$, $y = 16.3$, and thus $C = 16.3$ and

$$y = 16.3 + 1.064x - 0.0025x^2 \quad (\text{total cost})$$

$$\frac{y}{x} = \frac{16.3}{x} + 1.064 - 0.0025x \quad (\text{average cost})$$

EXAMPLE

Marginal cost y' as a function of units produced x is given by

$$y' = 2 + 60x - 5x^2$$

Find the total and average cost functions if fixed cost is 65.

$$y = \int (2 + 60x - 5x^2) dx$$

$$= 2x + 30x^2 - \frac{5}{3}x^3 + C$$

If $x = 0$, $y = 65$, and thus $C = 65$ and

$$y = 65 + 2x + 30x^2 - \frac{5}{3}x^3 \quad (\text{total cost})$$

$$\frac{y}{x} = \frac{65}{x} + 2 + 30x - \frac{5}{3}x^2 \quad (\text{average cost})$$

REVENUE

For any demand function

$$y = f(x)$$

where y is the price per unit and x is the number of units, *total revenue*, R , is the product of x and y ; that is,

$$R = xy = x \cdot f(x)$$

Marginal revenue with respect to demand is the derivative with respect to x of the total revenue

$$\frac{dR}{dx} = R'(x)$$

Thus the total revenue function is the integral with respect to x of the marginal revenue function; that is,

$$R = \int R'(x) dx$$

And, since

$$\int R'(x) dx = R(x) + C$$

an initial condition must be specified to obtain a unique total revenue function by integrating the corresponding marginal revenue function. The initial condition that revenue is zero if demand is zero can be used to evaluate the constant of integration.

Note that *average revenue* or revenue per unit is the price per unit, y , and thus the average revenue curve and the demand curve are identical.

EXAMPLE

If the marginal revenue function is

$$R'(x) = 8 - 6x - 2x^2$$

determine the total revenue and demand functions.

$$\begin{aligned} R(x) &= \int (8 - 6x - 2x^2) dx \\ &= 8x - 3x^2 - \frac{2}{3}x^3 + C \end{aligned}$$

If $x = 0$, $R = 0$, and thus $C = 0$ and

$$\begin{aligned} R &= 8x - 3x^2 - \frac{2}{3}x^3 \\ y = \frac{R}{x} &= 8 - 3x - \frac{2}{3}x^2 \end{aligned}$$

EXAMPLE

If the marginal revenue function is

$$R'(x) = 12 - 8x + x^2$$

determine the revenue and demand functions.

$$\begin{aligned} R(x) &= \int (12 - 8x + x^2) dx \\ &= 12x - 4x^2 + \frac{1}{3}x^3 + C \end{aligned}$$

If $x = 0$, $R = 0$, and thus $C = 0$ and

$$\begin{aligned} R &= 12x - 4x^2 + \frac{1}{3}x^3 \\ y = \frac{R}{x} &= 12 - 4x + \frac{1}{3}x^2 \\ &= \frac{(6 - x)^2}{3} \end{aligned}$$

NATIONAL INCOME, CONSUMPTION, AND SAVINGS

If the consumption function is given by

$$c = f(x)$$

where c is total national consumption and x is total national income, then the *marginal propensity to consume* is the derivative with respect to x of the consumption function

$$\frac{dc}{dx} = f'(x)$$

and, assuming $x = c + s$, where s is savings, the *marginal propensity to save* is

$$\frac{ds}{dx} = 1 - \frac{dc}{dx}$$

Total national consumption is the integral with respect to x of the marginal propensity to consume,

$$c = \int f'(x) dx = f(x) + C$$

An initial condition must be specified to obtain a unique consumption function by integrating the corresponding marginal propensity to consume.

EXAMPLE

The marginal propensity to consume (in billions of dollars) is

$$\frac{dc}{dx} = 0.7 + \frac{0.2}{\sqrt{x}}$$

When income is zero, consumption is 8 billion dollars. Find the consumption function.

$$\begin{aligned} c &= \int \left(0.7 + \frac{0.2}{\sqrt{x}} \right) dx \\ &= 0.7x + 0.4\sqrt{x} + C \end{aligned}$$

If $x = 0$, $c = 8$, and thus $C = 8$ and

$$c = 8 + 0.7x + 0.4\sqrt{x}$$

EXAMPLE

The marginal propensity to save is $\frac{1}{3}$. When income is zero, consumption is 11 billion dollars. Find the consumption function.

$$\frac{dc}{dx} = 1 - \frac{ds}{dx} = \frac{2}{3}$$

$$c = \int \frac{2}{3} dx$$

$$= \frac{2}{3}x + C$$

If $x = 0$, $c = 11$, and thus $C = 11$ and

$$c = \frac{2}{3}x + 11$$

CAPITAL FORMATION

The process of adding to a given stock of capital is referred to as *capital formation*. If this process is considered to be continuous over time, capital stock can be expressed as a function of time, $K(t)$, and the rate of capital formation is then given by $\frac{dK(t)}{dt} = K'(t)$. The rate of capital formation at time t is the same as the rate of *net investment* flow at time t , denoted by $I(t)$. Thus

$$K(t) = \int K'(t) dt = \int I(t) dt$$

and capital stock is the integral with respect to time of the rate of capital formation or, equivalently, the integral with respect to time of net investment. And, since

$$\int I(t) dt = K(t) + C$$

an initial condition must be specified to obtain a unique function for capital by integrating the rate of net-investment flow. For example, if the capital at time zero is a specified amount, $K(0)$, then this condition can be used to evaluate the constant of integration.

EXAMPLE

If investment flow is given by $I(t) = 5t^{3/7}$ and the initial capital stock at $t = 0$ is $K(0)$, find the function representing capital, K .

$$\begin{aligned} K &= \int 5t^{3/7} dt \\ &= \frac{7}{2}t^{10/7} + C \end{aligned}$$

If $t = 0$, $K = C$, and thus $C = K(0)$ and

$$K = \frac{7}{2}t^{10/7} + K(0)$$

EXAMPLE

If investment flow is given by $I(t) = 15t^{1/4}$ and the initial capital stock at $t = 0$ is 30, find the function representing capital, K .

$$\begin{aligned} K &= \int 15t^{1/4} dt \\ &= 12t^{5/4} + C \end{aligned}$$

If $t = 0$, $K = 30$, and thus $C = 30$ and

$$K = 12t^{5/4} + 30$$

PROBLEMS

1. If marginal revenue is a constant different from zero, show that price is constant.
2. If $R'(x) = 0$ and $R(0) \neq 0$, what is the nature of the demand curve?
3. If marginal cost is constant, show that the cost function is a straight line.
4. The marginal propensity to consume (in billions of dollars) is

$$\frac{dc}{dx} = 0.6 + \frac{0.5}{2x^{1/2}}$$

When income is zero, consumption is 10 billion dollars. Find the consumption function.

5. The marginal cost function for production is $y' = 10 + 24x - 3x^2$; if the (total) cost of producing 1 unit is 25, find the total cost function and the average cost function.
6. The marginal propensity to save is $\frac{1}{2}$. When income is zero, consumption is 6 billion dollars. Find the consumption function.
7. If marginal revenue is $R' = 15 - 9x - 3x^2$, find the revenue and demand functions.
8. If marginal revenue is $R' = \frac{3}{x^2} - \frac{2}{x}$, find the revenue and demand functions if $R(1) = 6$.

9. If marginal revenue is $R' = 10 - 5x$, find the revenue and demand functions.
 10. If marginal revenue is $R' = 20 - 3x^2$, find the revenue and demand functions.
 11. The marginal propensity to consume (in billions of dollars) is

$$\frac{dc}{dx} = 0.5 + \frac{1}{3x^{1/3}}$$

When income is zero, consumption is 6 billion dollars. Find the consumption function.

12. The marginal propensity to save (in billions of dollars) is

$$\frac{dc}{ds} = 1 - 0.4 - \frac{1}{6x^{2/3}}$$

When income is zero, consumption is 9 billion dollars. Find the consumption function.

13. If investment flow is given by $I(t) = 20t^{3/7}$ and the initial capital stock at $t = 0$ is 25, find the function representing capital, K .
 14. If investment flow is given by $I(t) = 25t^{4/11}$ and the initial capital stock at $t = 0$ is 22, find the function representing capital, K .

Answers to Odd-Numbered Problems

5. $y = 10x + 12x^2 - x^3 + 4$

9. $R = 10x - \frac{5}{2}x^2$

$\bar{y} = 10 + 12x - x^2 + \frac{4}{x}$

$y = 10 - \frac{5}{2}x$

7. $R = 15x - \frac{9}{2}x^2 - x^3$

11. $c = 0.5x + 0.5x^{2/3} + 6$

$y = 15 - \frac{9}{2}x - x^2$

13. $K = 14t^{10/7} + 25$

4.4 Definite Integration

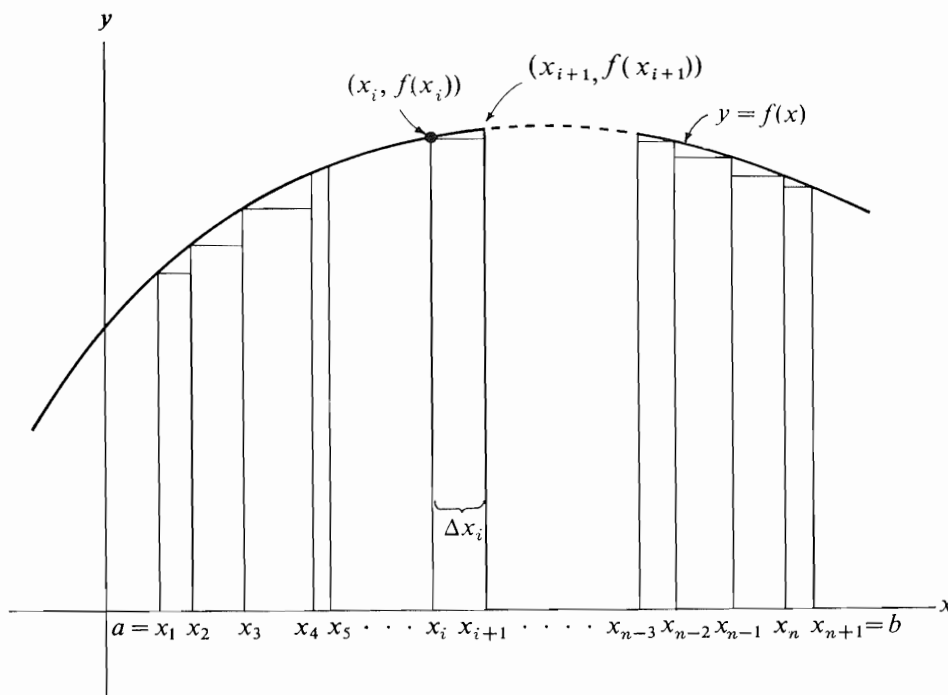
As noted above, finding the areas of figures of various shapes was one of the problems that led to the development of integral calculus. In elementary geometry, the area of a rectangle is shown to be equal to the product of its width and its height, and from this the areas of other figures bounded by straight-line segments are obtained by elementary geometrical methods. However, these methods are not directly applicable to figures bounded entirely or in part by curved lines. In general, to find areas of curvilinear figures, the method of limits must be used; for example, the area of a circle is obtained in geometry by considering it to be the common limit of sets of regular inscribed and circumscribed polygons as the number of their sides is indefinitely increased. This use of the method of limits leads to the interpretation of the definite integral as the area under a curve. In this section, the definite integral is defined and illustrated.

Consider the problem of determining the area bounded by the continuous positive curve $y = f(x)$, the x -axis, and the lines $x = a$ and $x = b$. Divide the base $[a, b]$ into n subintervals and denote the points of division by $a = x_1, x_2, \dots, x_n, x_{n+1} = b$ and the lengths of the n subintervals as $\Delta x_i = x_{i+1} - x_i; i = 1, \dots, n$. Erect ordinates at the points of division and inscribe rectangles (see Figure 4.3).

The areas of these inscribed rectangles are $f(x_1) \Delta x_1, f(x_2) \Delta x_2, \dots, f(x_n) \Delta x_n$ and their sum is

$$\sum_{i=1}^n f(x_i) \Delta x_i$$

Figure 4.3



By increasing the number of inscribed rectangles, that is, by letting $n \rightarrow \infty$ and $\max \Delta x_i \rightarrow 0$, the area under the curve between a and b which is not included in the rectangles is decreased and, in fact, approaches zero.

DEFINITION: The area bounded by the continuous positive function

$$y = f(x)$$

the x -axis, and two fixed ordinates $x = a$, $x = b$ is

$$A = \lim_{\substack{n \rightarrow \infty \\ \max \Delta x_i \rightarrow 0}} \sum_{i=1}^n f(x_i) \Delta x_i$$

NOTE: The same limit would be obtained by circumscribing rather than inscribing the rectangles, that is,

$$\lim_{\substack{n \rightarrow \infty \\ \max \Delta x_i \rightarrow 0}} \sum_{i=1}^n f(x_i) \Delta x_i = \lim_{\substack{n \rightarrow \infty \\ \max \Delta x_i \rightarrow 0}} \sum_{i=1}^n f(x_{i+1}) \Delta x_i$$

If $f(x)$ is nonmonotonic, the argument is modified by using, respectively, the minimum or maximum values of the function in the intervals rather than $f(x_i)$ or $f(x_{i+1})$. In fact, any point in an interval can be used without affecting the result.

If $f(x)$ is continuous for the interval $[a, b]$, then

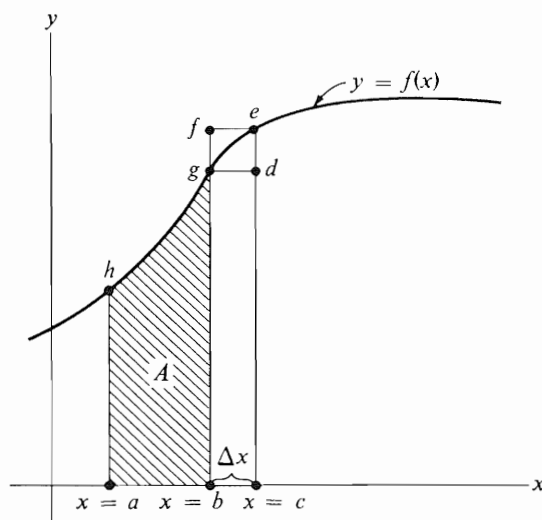
$$\lim_{\substack{n \rightarrow \infty \\ \max \Delta x_i \rightarrow 0}} \sum_{i=1}^n f(x_i) \Delta x_i$$

exists, $f(x)$ is said to be *integrable* over the interval $[a, b]$, and it can be shown that

$$\text{Area} = \lim_{\substack{n \rightarrow \infty \\ \max \Delta x_i \rightarrow 0}} \sum_{i=1}^n f(x_i) \Delta x_i = \int_a^b f(x) dx$$

by the following argument.

Figure 4.4



Consider the area between a fixed ordinate at $x = a$ and a variable (movable) ordinate at $x = b$ (see Figure 4.4) and denote the area $abgh$ by A . When x takes a small increment Δx , A takes an increment $\Delta A = \text{area } bceg$. Note that

$$\text{Area } (bcdg) < \text{Area } (bceg) < \text{Area } (bcef)$$

$$(bg)(\Delta x) < \Delta A < (ce)(\Delta x)$$

$$bg < \frac{\Delta A}{\Delta x} < ce$$

and let $\Delta x \rightarrow 0$. (If it happens that the shape of $f(x)$ is such that $ce < bg$, the inequality signs can be reversed without affecting the argument.) Then, since y is a continuous function of x , ce approaches bg and

$$\frac{dA}{dx} = bg = y = f(x)$$

$$dA = f(x) dx$$

Integrating,

$$A = \int f(x) dx$$

Denote $\int f(x) dx$ by $F(x) + C$. Then $A = F(x) + C$. Note that $A = 0$ when $x = a$ and thus

$$0 = F(a) + C$$

$$C = -F(a)$$

and

$$A = F(x) - F(a)$$

For the area $abgh$ (Figure 4.4), $x = b$ and

$$A = F(b) - F(a)$$

which is denoted by

$$\int_a^b y \, dx \quad \text{or} \quad \int_a^b f(x) \, dx$$

and is read “the integral from a to b of $y \, dx$.” This operation is referred to as *integration between limits*, a being the *lower limit* of integration and b being the *upper limit* of integration. In this context the word “limit” means the value of the variable at one end of its range; this should not be confused with the meaning of the word in the theory of limits. In the evaluation of $\int_a^b f(x) \, dx$, the constant of integration disappears; the integral thus has a definite value and, for this reason, is referred to as the *definite integral* of $f(x)$ from a to b . From the definition of a definite integral, it can be shown that definite integrals have the following properties.

1. $\int_a^b f(x) \, dx = -\int_b^a f(x) \, dx$.
2. $\int_a^a f(x) \, dx = 0$.
3. $\int_a^b f(x) \, dx = \int_a^c f(x) \, dx + \int_c^b f(x) \, dx$, where $a \leq c \leq b$.

NOTE: For some problems, definite integrals occur for which one or both limits of integration are infinite. An integral of the form

$$\int_{-\infty}^b f(x) \, dx \quad \int_a^{\infty} f(x) \, dx \quad \int_{-\infty}^{\infty} f(x) \, dx$$

is referred to as an *improper integral* and is evaluated by taking limits as follows:

$$\int_{-\infty}^b f(x) \, dx = \lim_{a \rightarrow -\infty} \int_a^b f(x) \, dx$$

$$\int_a^{\infty} f(x) \, dx = \lim_{b \rightarrow \infty} \int_a^b f(x) \, dx$$

$$\int_{-\infty}^{\infty} f(x) \, dx = \lim_{\substack{a \rightarrow -\infty \\ b \rightarrow \infty}} \int_a^b f(x) \, dx$$

If the limit exists, the improper integral *converges*; if the limit does not exist, the improper integral *diverges*.

EXAMPLE

(First Property)

$$\begin{aligned} \int_1^3 y(y^2 - 4)^2 \, dy &= \frac{1}{2} \int_1^3 (2y)(y^2 - 4)^2 \, dy \\ &= \left[\frac{(y^2 - 4)^3}{6} \right]_1^3 \\ &= \frac{12^3}{6} - \left(-\frac{2^3}{6} \right) \\ &= \frac{152}{3} = \frac{76}{3} \end{aligned}$$

$$\begin{aligned}
 \int_3^1 y(y^2 - 4)^2 dy &= \left[\frac{(y^2 - 4)^3}{6} \right]_3^1 \\
 &= -\frac{27}{6} - \frac{125}{6} \\
 &= -\frac{152}{6} = -\frac{76}{3}
 \end{aligned}$$

and

$$\int_1^3 y(y^2 - 4)^2 dy = -\int_3^1 y(y^2 - 4)^2 dy$$

(Second Property)

$$\begin{aligned}
 \int_{-1}^{-1} \sqrt{1-z} dz &= \left[-\frac{2}{3}(1-z)^{3/2} \right]_{-1}^{-1} \\
 &= -\frac{2}{3}(2)^{3/2} - \left[-\frac{2}{3}(2)^{3/2} \right] \\
 &= 0
 \end{aligned}$$

(Third Property)

$$\begin{aligned}
 \int_1^4 \frac{v+1}{\sqrt{v}} dv &= \int_1^4 (v^{1/2} + v^{-1/2}) dv \\
 &= \left[\frac{2}{3}v^{3/2} + 2v^{1/2} \right]_1^4 \\
 &= \frac{16}{3} + 4 - \frac{2}{3} - 2 \\
 &= \frac{20}{3} \\
 \int_1^4 \frac{v+1}{\sqrt{v}} dv &= \int_1^2 \frac{v+1}{\sqrt{v}} dv + \int_2^4 \frac{v+1}{\sqrt{v}} dv \\
 &= \left[\frac{2}{3}v^{3/2} + 2v^{1/2} \right]_1^2 + \left[\frac{2}{3}v^{3/2} + 2v^{1/2} \right]_2^4 \\
 &= \frac{2}{3}(2)^{3/2} + 2^{3/2} - \frac{2}{3} - 2 + \frac{16}{3} + 4 - \frac{2}{3}(2)^{3/2} - 2^{3/2} \\
 &= \frac{20}{3}
 \end{aligned}$$

PROBLEMS

Evaluate the following integrals.

1. $\int_0^1 (x^2 - 2x + 3) dx$

8. $\int_{-1}^1 (x^2 + 1)^2 dx$

2. $\int_{-1}^1 (v+1)^2 dv$

9. $\int_a^{2a} (a+z) dz$

3. $\int_0^2 (4x+1)^{1/2} dx$

10. $\int_1^2 \frac{x^2-1}{x^4} dx$

4. $\int_0^1 \frac{dx}{(2x+1)^3}$

11. $\int_1^8 (u^{1/3} - u^{-1/3}) du$

5. $\int_{-3}^{-2} t(t+1)^2 dt$

12. $\int_{-1}^0 (2x + x^2 - x^3) dx$

6. $\int_2^5 \left(x^2 + \frac{1}{x^2} \right) dx$

13. $\int_a^{2a} (a^3 + 3ax^2 + x^3) dx$

7. $\int_1^3 (2\theta + 1)(3 - \theta) d\theta$

14. $\int_0^a (\sqrt{a} - \sqrt{x})^2 dx$

15. $\int_1^4 (\sqrt{z} - z)^2 dz$

16. $\int_{-1}^2 (x^2 + x)(3x + 1) dx$

Answers to Odd-Numbered Problems

1. $\frac{7}{3}$

9. $\frac{7}{2}a^2$

3. $\frac{13}{3}$

11. $\frac{27}{4}$

5. $-\frac{73}{12}$

13. $\frac{47}{4}a^4$

7. $\frac{26}{3}$

15. $\frac{37}{10}$

AREA AS A DEFINITE INTEGRAL

The previous result concerning evaluation of area by integration is known as the *fundamental theorem of integral calculus* and can be summarized in somewhat more rigorous form as follows:

Let $f(x)$ be continuous and positive for the interval $x = a$ to $x = b$. Let this interval be divided into n subintervals of lengths $\Delta x_1, \Delta x_2, \dots, \Delta x_n$ and points be chosen, one in each subinterval, their abscissas being x_1, x_2, \dots, x_n , respectively. Then

$$\lim_{\substack{n \rightarrow \infty \\ \max \Delta x_i \rightarrow 0}} \sum_{i=1}^n f(x_i) \Delta x_i = \int_a^b f(x) dx = F(b) - F(a)$$

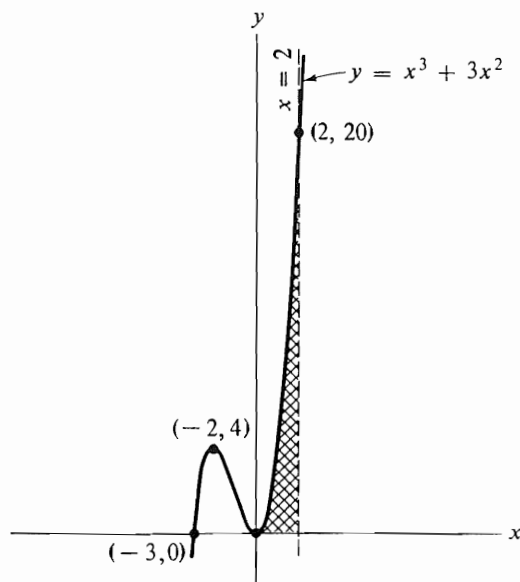
The importance of this theorem results from the fact that it permits the evaluation of the limit of a sum of terms by integration and that this limit can be interpreted as the area under a curve. Specifically, the definite integral $\int_a^b f(x) dx$ can be interpreted as the area bounded by the continuous positive curve $y = f(x)$, the x -axis, and the lines $x = a$ and $x = b$.

EXAMPLE

Find the area bounded by the curve

$$y = x^3 + 3x^2$$

Figure 4.5



the x -axis, and the lines $x = 0$ and $x = 2$ (see Figure 4.5).

$$A = \int_0^2 (x^3 + 3x^2) dx = \left[\frac{x^4}{4} + x^3 \right]_0^2 = (4 + 8) - 0 = 12$$

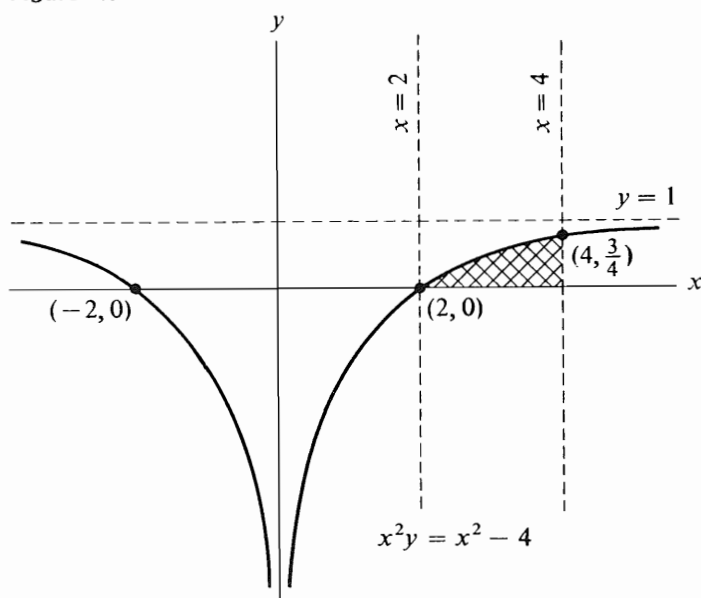
EXAMPLE

Find the area bounded by the curve

$$x^2y = x^2 - 4$$

the x -axis, and the lines $x = 2$ and $x = 4$ (see Figure 4.6).

Figure 4.6



$$\begin{aligned} A &= \int_2^4 \frac{x^2 - 4}{x^2} dx = \int_2^4 \left(1 - \frac{4}{x^2} \right) dx = [x + 4x^{-1}]_2^4 \\ &= (4 + 1) - (2 + 2) \\ &= 1 \end{aligned}$$

EXAMPLE

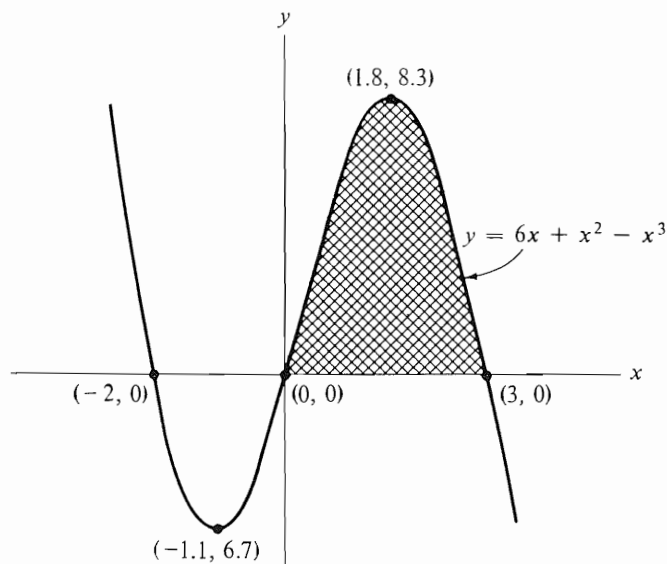
Find the area in the first quadrant bounded by the x -axis and the curve

$$y = 6x + x^2 - x^3$$

(see Figure 4.7).

$$A = \int_0^3 (6x + x^2 - x^3) dx = \left[3x^2 + \frac{x^3}{3} - \frac{x^4}{4} \right]_0^3 = (27 + 9 - \frac{81}{4}) - 0 = 15\frac{3}{4}$$

Figure 4.7

**EXAMPLE**

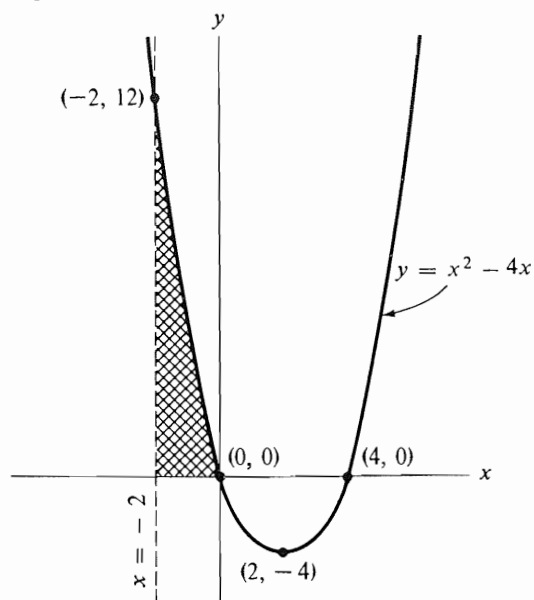
Find the total area between the parabola

$$y = x^2 - 4x$$

the x -axis, and the line $x = -2$ (see Figure 4.8).

$$A = \int_{-2}^0 (x^2 - 4x) dx = \left[\frac{x^3}{3} - 2x^2 \right]_{-2}^0 = 0 - \left(-\frac{8}{3} - 8 \right) = 10\frac{2}{3}$$

Figure 4.8



Interpretation of Negative Areas

In the definition of area given above,

$$A = \int_a^b f(x) dx$$

$f(x)$ is assumed to be a continuous positive function between a and b . If $f(x)$ is negative between a and b , that is, if the curve $y = f(x)$ lies below the x -axis between a and b , then the value of the integral

$$A = \int_a^b f(x) dx$$

is negative. Such areas below the x -axis are called *negative areas*; the total absolute area between a curve, the x -axis, and two ordinates is given by

$$\text{Total area} = \sum (\text{positive areas}) - \sum (\text{negative areas})$$

NOTE: This is equivalent to saying that area is equal to the absolute value of the integral and thus is always positive.

EXAMPLE

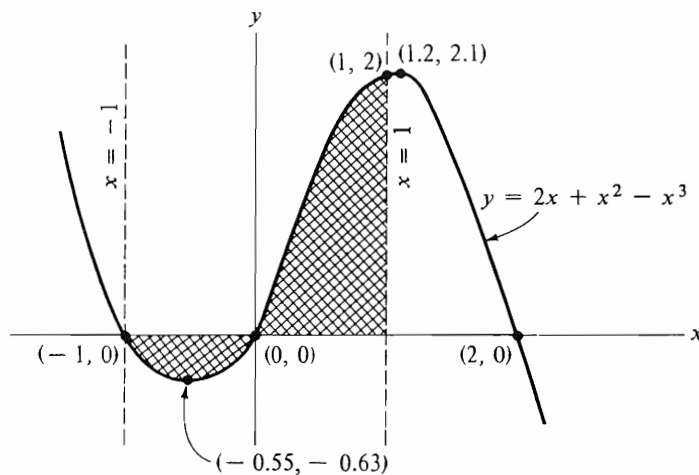
Find the area bounded by the curve

$$y = 2x + x^2 - x^3$$

the x -axis, and the lines $x = -1$ and $x = 1$ (see Figure 4.9).

$$\begin{aligned} A &= \int_0^1 (2x + x^2 - x^3) dx - \int_{-1}^0 (2x + x^2 - x^3) dx \\ &= \left[x^2 + \frac{x^3}{3} - \frac{x^4}{4} \right]_0^1 - \left[x^2 + \frac{x^3}{3} - \frac{x^4}{4} \right]_{-1}^0 \\ &= \left[\left(1 + \frac{1}{3} - \frac{1}{4} \right) - 0 \right] - \left[0 - \left(1 - \frac{1}{3} - \frac{1}{4} \right) \right] \\ &= \frac{13}{12} - \left(-\frac{5}{12} \right) \text{ (positive area minus negative area)} \\ &= \frac{3}{2} \end{aligned}$$

Figure 4.9



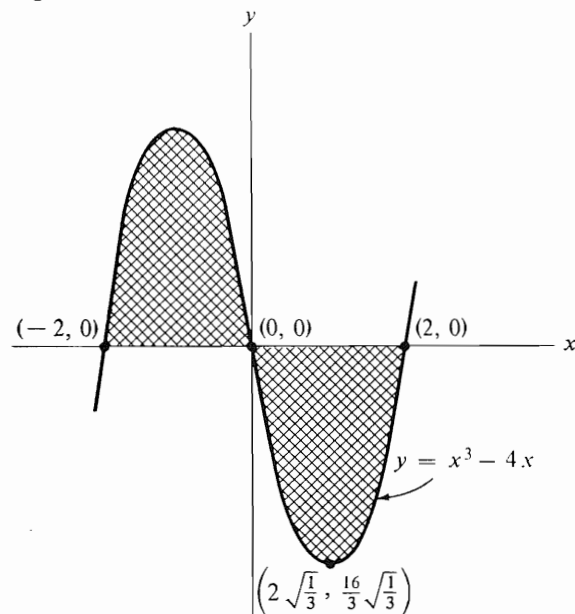
EXAMPLE

Find the area bounded by the curve

$$y = x^3 - 4x$$

and the x -axis (see Figure 4.10).

Figure 4.10



$$\begin{aligned}
 A &= \int_{-2}^0 (x^3 - 4x) \, dx - \int_0^2 (x^3 - 4x) \, dx \\
 &= \left[\frac{x^4}{4} - 2x^2 \right]_{-2}^0 - \left[\frac{x^4}{4} - 2x^2 \right]_0^2 \\
 &= [0 - (4 - 8)] - [(4 - 8) - 0] \\
 &= 4 - (-4) \\
 &= 8
 \end{aligned}$$

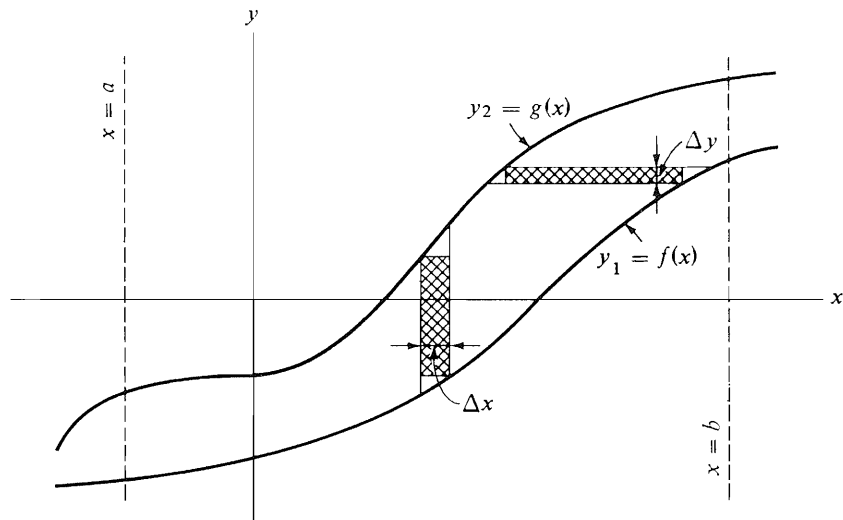
Area Between Two Curves

Suppose that the area to be evaluated is between the curves $y_1 = f(x)$ and $y_2 = g(x)$ and the lines $x = a$ and $x = b$ and that (for definiteness) $f(x) \leq g(x)$ for $a \leq x \leq b$ (see Figure 4.11). Then

$$A = \int_a^b [g(x) - f(x)] \, dx$$

Note that this formula includes negative areas (with appropriate signs) in the total area between the curves.

Figure 4.11



EXAMPLE

Find the area bounded by the curves

$$y = x^2 \quad \text{and} \quad y = x$$

(see Figure 4.12).

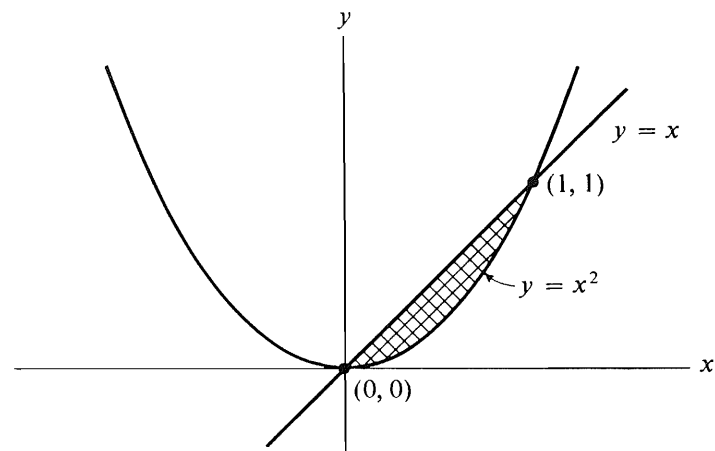
Finding the points of intersection of the curves,

$$y = x = x^2$$

$$x(1 - x) = 0$$

$$x = 0, 1$$

Figure 4.12



If $x = 0$, $y = 0$; if $x = 1$, $y = 1$. Thus

$$\begin{aligned} A &= \int_0^1 (x - x^2) dx = \left[\frac{x^2}{2} - \frac{x^3}{3} \right]_0^1 \\ &= \left(\frac{1}{2} - \frac{1}{3} \right) - 0 \\ &= \frac{1}{6} \end{aligned}$$

EXAMPLE

Find the area bounded by the curves

$$y = x^3 \quad \text{and} \quad y = 2x^2$$

(see Figure 4.13).

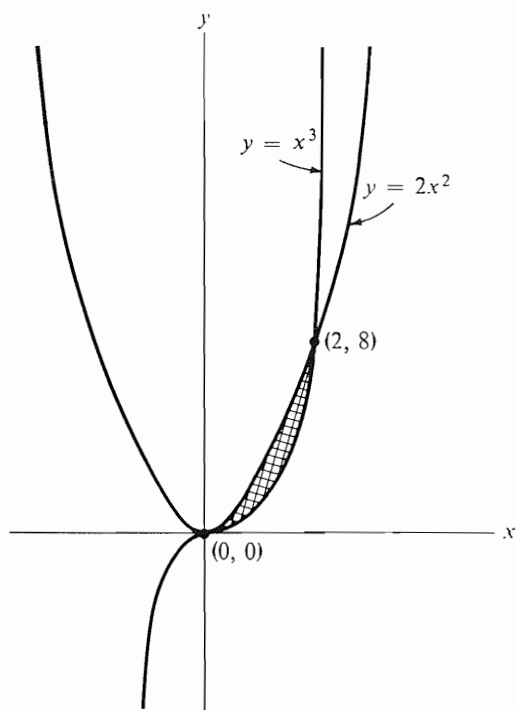
Finding the points of intersection of the curves,

$$\begin{aligned} y &= x^3 = 2x^2 \\ x^2(x - 2) &= 0 \\ x &= 0, 2 \end{aligned}$$

If $x = 0$, $y = 0$; if $x = 2$, $y = 8$. Thus

$$\begin{aligned} A &= \int_0^2 (2x^2 - x^3) dx = \left[\frac{2x^3}{3} - \frac{x^4}{4} \right]_0^2 \\ &= \left(\frac{16}{3} - 4 \right) - 0 \\ &= \frac{4}{3} \end{aligned}$$

Figure 4.13



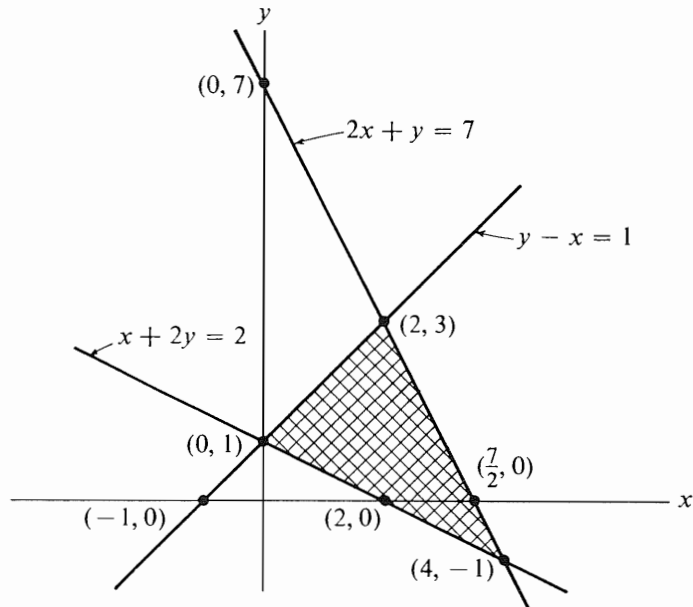
EXAMPLE

Find the area bounded by the curves

$$x + 2y = 2 \quad y - x = 1 \quad \text{and} \quad 2x + y = 7$$

(see Figure 4.14).

Figure 4.14



Finding the points of intersection of the straight lines,

$$x + 2y = 2 \quad x + 2y = 2 \quad y - x = 1$$

$$y - x = 1 \quad 2x + y = 7 \quad 2x + y = 7$$

$$3y = 3 \quad -3y = 3 \quad 3y = 9$$

$$y = 1 \quad y = -1 \quad y = 3$$

$$x = 0 \quad x = 4 \quad x = 2$$

$$\begin{aligned} A &= \int_0^2 \left[(1+x) - \left(\frac{2-x}{2} \right) \right] dx + \int_2^4 \left[(7-2x) - \left(\frac{2-x}{2} \right) \right] dx \\ &= \int_0^2 \frac{3}{2}x \, dx + \int_2^4 \left(6 - \frac{3}{2}x \right) dx = \left[\frac{3x^2}{4} \right]_0^2 + \left[6x - \frac{3x^2}{4} \right]_2^4 \\ &= 3 - 0 + (24 - 12) - (12 - 3) \\ &= 6 \end{aligned}$$

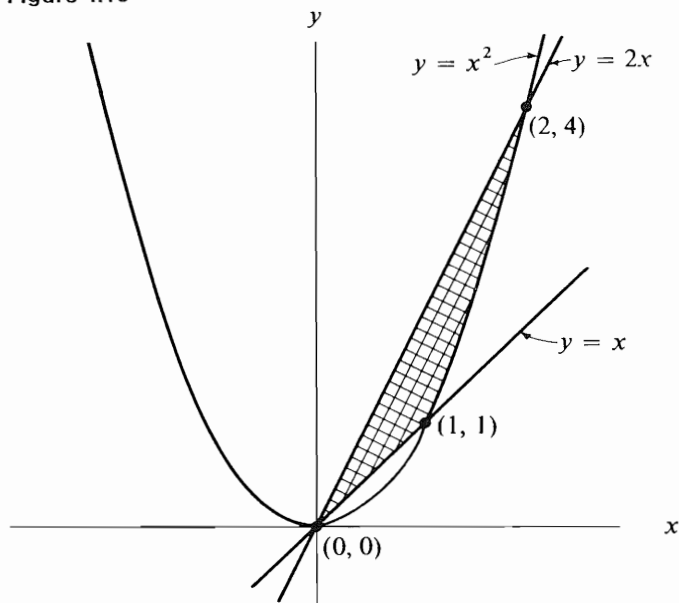
EXAMPLE

Find the area bounded by the curves

$$y = x^2 \quad y = x \quad y = 2x$$

(see Figure 4.15).

Figure 4.15



Finding the points of intersection of the curves,

$$\begin{array}{lll}
 y = x = x^2 & y = 2x = x^2 & y = x = 2x \\
 x(x-1) = 0 & x(x-2) = 0 & x = 0 \\
 x = 0, 1 & x = 0, 2 & y = 0
 \end{array}$$

$$\text{If } x = 0, y = 0 \quad \text{If } x = 0, y = 0$$

$$\text{If } x = 1, y = 1 \quad \text{If } x = 2, y = 4$$

Thus

$$\begin{aligned}
 A &= \int_0^1 (2x - x) dx + \int_1^2 (2x - x^2) dx \\
 &= \left[\frac{x^2}{2} \right]_0^1 + \left[x^2 - \frac{x^3}{3} \right]_1^2 \\
 &= \frac{1}{2} - 0 + \left(4 - \frac{8}{3} \right) - \left(1 - \frac{1}{3} \right) \\
 &= \frac{7}{6}
 \end{aligned}$$

NOTE: In evaluating an area it is theoretically of no importance whether horizontal or vertical elements of area are used (see Figure 4.11); the choice depends on the number and difficulty of the integrals needed to determine the particular area. In the above examples, the curves are easily written in the form $x = g(y)$ and horizontal elements of area can conveniently be used.

EXAMPLE

The area bounded by the curves

$$y = x^2 \quad \text{and} \quad y = x$$

could have been determined using horizontal elements of area as follows:

$$A = \int_0^1 (y^{1/2} - y) dy = \left[\frac{2y^{3/2}}{3} - \frac{y^2}{2} \right]_0^1 = \left(\frac{2}{3} - \frac{1}{2} \right) - 0 = \frac{1}{6}$$

EXAMPLE

The area bounded by the curves

$$y = x^3 \quad \text{and} \quad y = 2x^2$$

could have been determined using horizontal elements of area as follows:

$$\begin{aligned} A &= \int_0^8 \left(y^{1/3} - \frac{y^{1/2}}{2^{1/2}} \right) dy = \left[\frac{3y^{4/3}}{4} - \frac{2^{1/2}y^{3/2}}{3} \right]_0^8 \\ &= \left(12 - \frac{32}{3} \right) - 0 = \frac{4}{3} \end{aligned}$$

EXAMPLE

The area bounded by the curves

$$x + 2y = 2 \quad y - x = 1 \quad 2x + y = 7$$

could have been determined using horizontal elements of area as follows:

$$\begin{aligned} A &= \int_{-1}^1 \left[\left(\frac{7-y}{2} \right) - (2-2y) \right] dy + \int_1^3 \left[\left(\frac{7-y}{2} \right) - (y-1) \right] dy \\ &= \int_{-1}^1 \frac{3}{2}(y+1) dy + \int_1^3 \frac{3}{2}(3-y) dy \\ &= \left[\frac{3y^2}{4} + \frac{3y}{2} \right]_{-1}^1 + \left[\frac{9y}{2} - \frac{3y^2}{4} \right]_1^3 \\ &= \left(\frac{3}{4} + \frac{3}{2} \right) - \left(\frac{3}{4} - \frac{3}{2} \right) + \left(\frac{27}{2} - \frac{27}{4} \right) - \left(\frac{9}{2} - \frac{3}{4} \right) \\ &= 6 \end{aligned}$$

EXAMPLE

The area bounded by the curves

$$y = x^2 \quad y = x \quad y = 2x$$

could have been determined using horizontal elements of area as follows:

$$\begin{aligned} A &= \int_0^1 \left(y - \frac{y}{2} \right) dy + \int_1^4 \left(y^{1/2} - \frac{y}{2} \right) dy \\ &= \left[\frac{y^2}{4} \right]_0^1 + \left[\frac{2y^{3/2}}{3} - \frac{y^2}{4} \right]_1^4 \\ &= \frac{1}{4} - 0 + \left(\frac{16}{3} - 4 \right) - \left(\frac{2}{3} - \frac{1}{4} \right) \\ &= \frac{7}{6} \end{aligned}$$

Probability as an Area

One of the most useful applications of definite integration to obtain the area under a curve is in evaluating probabilities. The probability of an event can be obtained by determining the corresponding area under a frequency function. A *frequency function* for a continuous variable x is a function $f(x)$ having the following properties

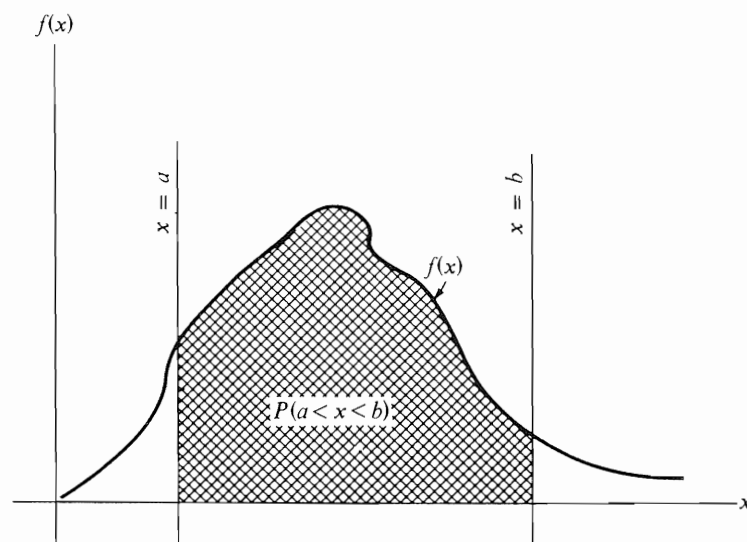
1. $f(x) \geq 0$
2. $\int_{-\infty}^{\infty} f(x) dx = 1$
3. $\int_a^b f(x) dx = P(a < x < b)$

where a and b are any two values of x and $a < b$.

These properties state that (1) probability is always nonnegative; (2) the probability of an event that is certain to occur is one; and (3) the probability that the value of x is in the interval (a, b) is equal to the area bounded by $y = f(x)$, $y = 0$, $x = a$, and $x = b$. (See Figure 4.16).

NOTE: If $f(x)$ is continuous, $P(a < x < b) = P(a \leq x \leq b)$.

Figure 4.16



EXAMPLE

In an autoparts warehouse, the proportion of orders filled per day has a frequency function given by

$$f(x) = 20(x^3 - x^4) \quad 0 \leq x \leq 1$$

- (a) What is the probability that less than 20 percent of the orders will be filled in a day?
- (b) What is the probability that between 90 and 100 percent of the orders will be filled in a day?

$$\begin{aligned}
 \text{(a) } P(x < 0.20) &= 20 \int_0^{0.2} (x^3 - x^4) dx \\
 &= 20 \left[\frac{x^4}{4} - \frac{x^5}{5} \right]_0^{0.2} \\
 &= 20 \left[\left(\frac{0.0016}{4} \right) - \left(\frac{0.00032}{5} \right) - 0 \right] \\
 &= 0.008 - 0.00128 \\
 &= 0.00672
 \end{aligned}$$

(See area *a* in Figure 4.17.)

Thus the probability of filling less than 20 percent of the orders is 0.00672, which means that this is quite unlikely to occur.

$$\begin{aligned}
 \text{(b) } P(0.9 \leq x \leq 1) &= 20 \int_{0.9}^1 (x^3 - x^4) dx \\
 &= 20 \left[\frac{x^4}{4} - \frac{x^5}{5} \right]_{0.9}^1 \\
 &= 20 \left[\left(\frac{1}{4} - \frac{1}{5} \right) - \left(\frac{0.6561}{4} - \frac{0.59049}{5} \right) \right] \\
 &= (5 - 4) - (3.2805 - 2.36196) \\
 &= 0.08146
 \end{aligned}$$

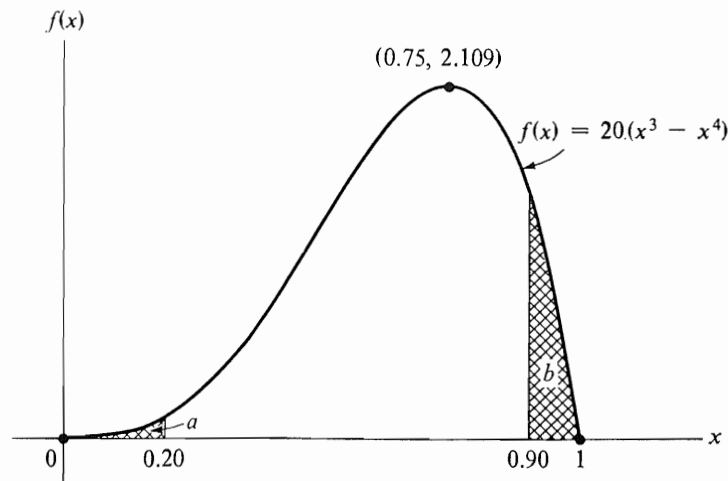
(See area *b* in Figure 4.17.)

Thus the probability of filling between 90 and 100 percent of the orders is 0.08146.

Note that, since $f(x) = 20(x^3 - x^4)$ is defined only for $0 \leq x \leq 1$ and is zero elsewhere,

$$\begin{aligned}
 \int_{-\infty}^{\infty} 20(x^3 - x^4) dx &= \int_0^1 20(x^3 - x^4) dx \\
 &= 20 \left[\frac{x^4}{4} - \frac{x^5}{5} \right]_0^1 \\
 &= 1
 \end{aligned}$$

Figure 4.17



So $\int_{-\infty}^{\infty} f(x) dx = 1$ as required for a frequency function; note also that $f(x) \geq 0$ for $0 \leq x \leq 1$ and is zero elsewhere.

EXAMPLE

Suppose that measurements of daily temperature are recorded to the nearest tenth of a degree. Then the difference between the true and recorded temperatures, the rounding error, is between -0.05 and 0.05 degrees. If the error is uniformly distributed in this interval, its frequency function is represented by

$$f(x) = \frac{1}{0.1} = 10 \quad -0.05 \leq x \leq 0.05$$

Note that, in general, if x is uniformly distributed on the interval $a \leq x \leq b$, then

$$f(x) = \frac{1}{b-a} \quad a \leq x \leq b$$

and is zero elsewhere.

- (a) What is the probability that the rounding error is between -0.01 and 0.01 ?
 (b) What is the probability that the rounding error is between 0.04 and 0.05 ?

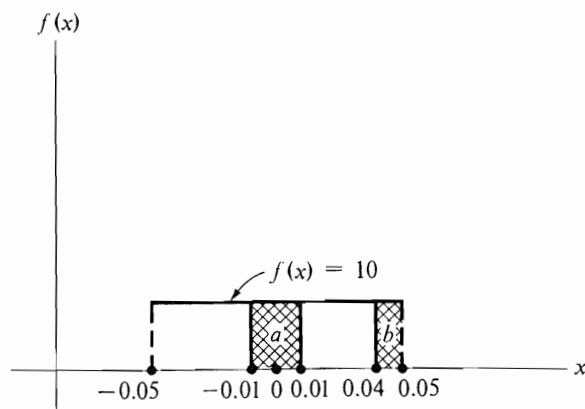
$$\begin{aligned} \text{(a) } P(-0.01 \leq x \leq 0.01) &= \int_{-0.01}^{0.01} 10 dx \\ &= [10x]_{-0.01}^{0.01} \\ &= 0.1 - (-0.1) = 0.2 \end{aligned}$$

(See area a in Figure 4.18.)

$$\begin{aligned} \text{(b) } P(0.04 \leq x \leq 0.05) &= \int_{0.04}^{0.05} 10 dx \\ &= [10x]_{0.04}^{0.05} \\ &= 0.5 - 0.4 = 0.1 \end{aligned}$$

(See area b in Figure 4.18.)

Figure 4.18



Note that

$$\begin{aligned}\int_{-\infty}^{\infty} 10 \, dx &= \int_{-0.05}^{0.05} 10 \, dx = [10x]_{-0.05}^{0.05} \\ &= 0.5 - (-0.5) = 1\end{aligned}$$

So $\int_{-\infty}^{\infty} f(x) \, dx = 1$ as required for a frequency function; note also that $f(x) \geq 0$ for $-0.05 \leq x \leq 0.05$ and is zero elsewhere.

PROBLEMS

Sketch each of the following curves and find the area bounded by the curve, the x -axis, and the given ordinates.

1. $y = \sqrt{x}$; $x = 1$, $x = 16$
2. $y = 2x + 1$; $x = 0$, $x = 4$
3. $y = x^{99}$; $x = 0$, $x = 1$
4. $y = 3x^2$; $x = 1$, $x = 3$
5. $y = x^2 - 3x$; $x = -1$, $x = 4$
6. $y = -x^2 + 4x$ (and x -axis)
7. $f(x) = \begin{cases} x^2 & x \leq 2 \\ -x + 6 & x > 2 \end{cases} \quad x = 0, x = 3$
8. $f(x) = \begin{cases} 2x + 3 & x \leq 3 \\ -x + 12 & x > 3 \end{cases} \quad x = 2, x = 5$
9. Find the area between the curve $y = 2x^4 - x^2$, the x -axis, and the two minimum ordinates.
10. Find the area bounded by the coordinate axes and the parabolic arc $\sqrt{x} + \sqrt{y} = \sqrt{a}$.

Draw a sketch and find the area bounded by the following curves.

- | | |
|--|---|
| 11. $y^2 = 2x$
$y = x - 4$ | 21. $y = x^2$
$y = x + 2$ |
| 12. $x^2 = 2ay$
$y = 2a$ | 22. $y = 4x - 4$
$y = \frac{1}{3}x^2$ |
| 13. $y = x - x^2$
$y = -x$ | 23. $y = x^3 + 3x^2 + 2$
$y = x^3 + 6x^2 - 25$ |
| 14. $y^2 = 4ax$
$x^2 = 4ay$ | 24. $y = 25 - x^2$
$y = (5 - x)^2$ |
| 15. $y^2 = x$
$y = x^3$ | 25. $y = x^3 - 3x^2 - 10x$
$y = -6x$ |
| 16. $y = (x - 1)^3$
$y = x^2 - x - 1$ | 26. $y = (x + 2)(x - 1)(x - 5)$
$y = (x + 2)(x - 1)$ |
| 17. $y^2 = 5a^2 - ax$
$y^2 = 4ax$ | 27. $y = x(x - 3)(x + 3)$
$y = -5x$ |
| 18. $x^2y = 4$
$y = 7 - 3x$ | 28. $y = x^3 + 3x^2 + 6$
$y = x^3 + 4x^2 + 5x$ |
| 19. $y = x^2$
$y = 8 - x^2$
$y = 4x + 12$ | 29. $y = x^3 - 5x^2 - 8x + 12$
$y = x^3 - 6x^2 + 21$ |
| 20. $y^3 = x^2$
$2x + y + 1 = 0$
$x - y = 4$ | |

HINTS

15. $y^2 = x$ and $y = x^3$ intersect $(0, 0)$ and $(1, 1)$
18. $x^2y = 4$ and $y = 7 - 3x$ intersect $(1, 4)$ and $(2, 1)$
20. $y^3 = x^2$ and $2x + y + 1 = 0$ intersect $(-1, 1)$
 $y^3 = x^2$ and $x - y = 4$ intersect $(8, 4)$

29. $x^3 - 5x^2 - 8x + 12 = (x + 2)(x - 1)(x - 6)$
 30. A computer center is open 12 hours a day and repairs, except for emergencies, are scheduled for the other 12 hours. The frequency function for the number of hours the center is actually in operation is given by

$$f(x) = \frac{1}{24} + \frac{x^2}{1152} \quad 0 \leq x \leq 12$$

- a. Find the probability that the center is in operation between 10 and 12 hours a day. b. Find the probability that the center is in operation less than 6 hours a day.
 31. A firm has a very large number of company cars for use by its employees. Records of the time each car is out of service for repairs are kept as the basis for deciding when a car should be sold. The frequency function for the total number of days a car is out of service before it is considered too expensive to repair and is sold is given by

$$f(x) = 0.2e^{-0.2x} \quad 0 \leq x < \infty$$

- a. Find the probability that a car is out of service a total of more than 30 days before being sold. b. Find the probability that a car is out of service a total of less than 5 days before being sold.
 32. Determine whether each of the following is a frequency function.

a. $f(x) = x^{-2}, \quad 1 \leq x < \infty$

d. $f(x) = \frac{1}{10}, \quad 1 \leq x \leq 20$

b. $f(x) = \frac{3}{8} - \frac{x^2}{576}, \quad 0 \leq x \leq 24$

e. $f(x) = \frac{4x^3}{15}, \quad 1 \leq x \leq 2$

c. $f(x) = 2e^{-4x}, \quad 0 \leq x < \infty$

f. $f(x) = \frac{144}{5}x - 6, \quad \frac{1}{3} \leq x \leq \frac{1}{2}$

Answers to Odd-Numbered Problems

- | | |
|-----------------------|----------------------|
| 1. 42 | 19. 64 |
| 3. 0.01 | 21. $\frac{23}{6}$ |
| 5. $\frac{49}{6}$ | 23. 108 |
| 7. $\frac{37}{6}$ | 25. $160\frac{3}{4}$ |
| 9. $\frac{7}{120}$ | 27. 8 |
| 11. 18 | 29. $166\frac{2}{3}$ |
| 13. $\frac{4}{3}$ | 31. a. 0.002 |
| 15. $\frac{5}{12}$ | b. 0.632 |
| 17. $\frac{40}{3}a^2$ | |

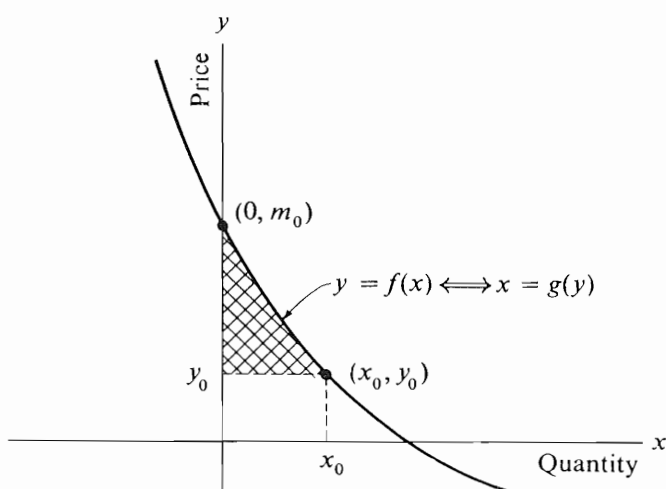
4.5 Applications of Definite Integration in Business and Economics

Definite integration has a variety of applications in business and economics; the concepts of *consumer's surplus* and *producers' surplus* are two examples of such applications. Other examples involve the determination of the revenue produced for a firm by a factor of production over a period of time in relation to the cost of that factor over the same period of time; in such examples the point in time at which revenue produced equals cost of the factor can also be determined. Applications involving consumers' surplus, producers' surplus, and revenue vs. cost are discussed in this section.

CONSUMERS' SURPLUS

A demand function represents the quantities of a commodity that would be purchased at various prices. If the market price is y_0 and the corresponding market demand is x_0 , then those consumers who would be willing to pay more than this

Figure 4.19



market price gain from the fact that the price is only y_0 (see Figure 4.19). Under certain economic assumptions, the total consumer gain is represented by the area below the demand curve and above the line $y = y_0$. This area is designated by Marshall as *consumers' surplus* and is evaluated as

$$\text{consumers' surplus} = \int_0^{x_0} f(x) \, dx - x_0 y_0$$

where the demand function is $y = f(x)$, or as

$$\text{consumers' surplus} = \int_{y_0}^{m_0} g(y) \, dy$$

where the demand function is $x = g(y)$ and m_0 is the value of y when $x = 0$, that is, m_0 is the y -intercept of the demand function. Thus

$$\text{consumers' surplus} = \int_0^{x_0} f(x) \, dx - x_0 y_0 = \int_{y_0}^{m_0} g(y) \, dy$$

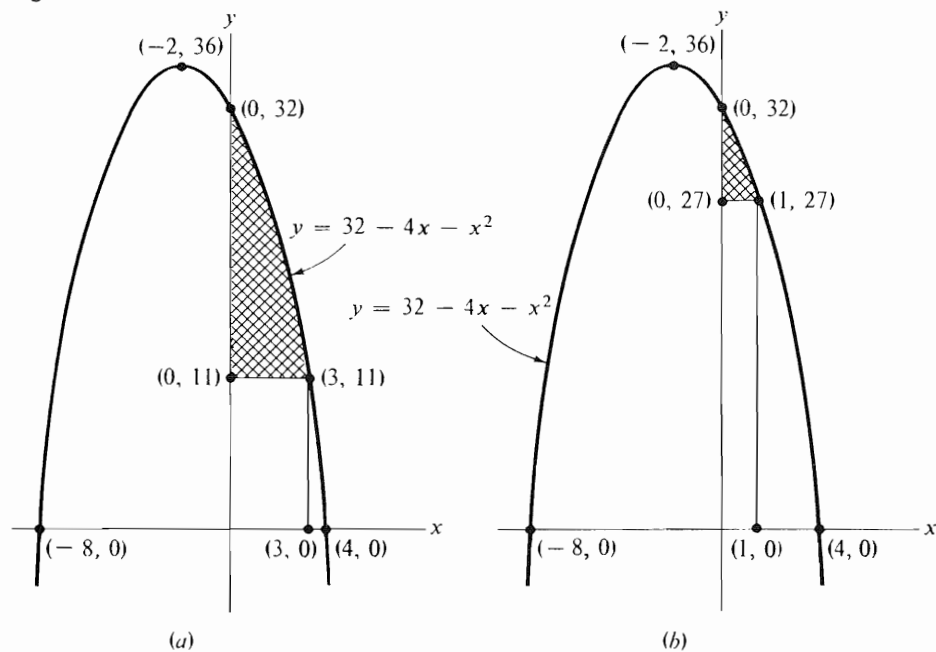
EXAMPLE

If the demand function is $y = 32 - 4x - x^2$, find the consumers' surplus

- (a) if $x_0 = 3$,
 (b) if $y_0 = 27$. (See Figure 4.20.)

$$\begin{aligned} \text{(a) consumers' surplus} &= \int_0^3 (32 - 4x - x^2) \, dx - (3)(11) \\ &= \left[32x - 2x^2 - \frac{x^3}{3} \right]_0^3 - 33 \\ &= (96 - 18 - 9) - 0 - 33 \\ &= 36 \end{aligned}$$

Figure 4.20



$$\begin{aligned}
 \text{(b) consumers' surplus} &= \int_0^1 (32 - 4x - x^2) dx - (1)(27) \\
 &= \left[32x - 2x^2 - \frac{x^3}{3} \right]_0^1 - 27 \\
 &= (32 - 2 - \frac{1}{3}) - 0 - 27 \\
 &= \frac{8}{3}
 \end{aligned}$$

EXAMPLE

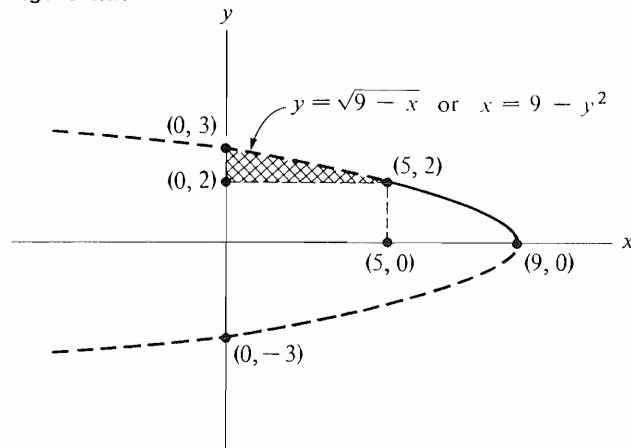
If the demand function is $y = \sqrt{9 - x}$ and $x_0 = 5$, find the consumers' surplus using two different methods (see Figure 4.21).

$$\begin{aligned}
 \text{consumers' surplus} &= \int_0^5 (9 - x)^{1/2} dx - (5)(2) = \int_2^3 (9 - y^2) dy \\
 \int_0^5 (9 - x)^{1/2} dx - 10 &= \left[-\frac{2}{3}(9 - x)^{3/2} \right]_0^5 - 10 \\
 &= -\frac{16}{3} - (-18) - 10 = \frac{8}{3}
 \end{aligned}$$

Alternatively,

$$\begin{aligned}
 \int_2^3 (9 - y^2) dy &= \left[9y - \frac{y^3}{3} \right]_2^3 \\
 &= (27 - 9) - (18 - \frac{8}{3}) = \frac{8}{3} \quad (\text{as above})
 \end{aligned}$$

Figure 4.21



EXAMPLE

The quantity sold and the corresponding price, under a monopoly, are determined by the demand function $y = 16 - x^2$ and by the marginal cost $y' = 6 + x$ in such a way as to maximize profit. Determine the corresponding consumers' surplus (see Figure 4.22).

$$\text{revenue} = 16x - x^3$$

$$\text{marginal revenue} = 16 - 3x^2$$

Profit is maximized when marginal revenue equals marginal cost—that is, when

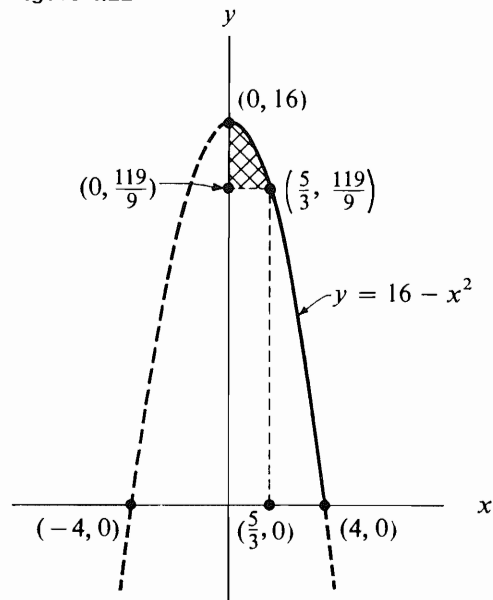
$$16 - 3x^2 = 6 + x$$

$$3x^2 + x - 10 = 0$$

$$(3x - 5)(x + 2) = 0$$

$$x = \frac{5}{3}, -2$$

Figure 4.22



so $x_0 = \frac{5}{3}$, $y_0 = \frac{119}{9}$.

$$\begin{aligned} \text{consumers' surplus} &= \int_0^{5/3} (16 - x^2) dx - \left(\frac{5}{3}\right)\left(\frac{119}{9}\right) \\ &= \left[16x - \frac{x^3}{3}\right]_0^{5/3} - \frac{595}{27} \\ &= \left(\frac{80}{3} - \frac{125}{81}\right) - 0 - \frac{595}{27} \\ &= \frac{250}{81} \approx 3.09 \end{aligned}$$

PRODUCERS' SURPLUS

A supply function represents the respective quantities of a commodity that would be supplied at various prices. If the market price is y_0 and the corresponding market supply is x_0 , then those producers who would be willing to supply the commodity below this market price gain from the fact that the price is y_0 . Under certain economic assumptions the total producer gain is represented by the area above the supply curve and below the line $y = y_0$ and is known as *producers' surplus* (see Figure 4.23). This area is evaluated as

$$\text{producers' surplus} = x_0 y_0 - \int_0^{x_0} f(x) dx$$

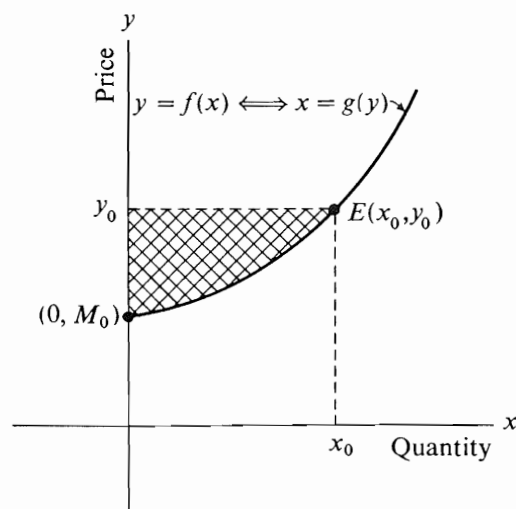
where the supply function is $y = f(x)$ and as

$$\text{producers' surplus} = \int_{M_0}^{y_0} g(y) dy$$

where the supply function is $x = g(y)$ and M_0 is the value of y when $x = 0$, that is, M_0 is the y -intercept of the supply function.

$$\text{producers' surplus} = x_0 y_0 - \int_0^{x_0} f(x) dx = \int_{M_0}^{y_0} g(y) dy$$

Figure 4.23



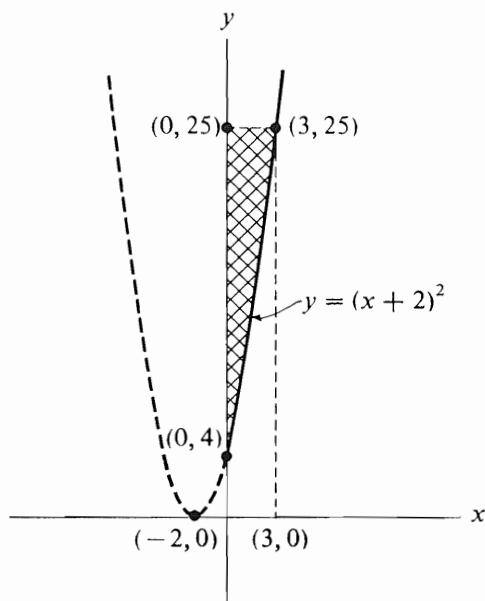
EXAMPLE

If the supply function is $y = (x + 2)^2$ and the price is $y_0 = 25$, find the producers' surplus using two different methods (see Figure 4.24).

$$\text{producers' surplus} = (3)(25) - \int_0^3 (x + 2)^2 dx = \int_4^{25} (y^{1/2} - 2) dy$$

$$75 - \int_0^3 (x + 2)^2 dx = 75 - \left[\frac{(x + 2)^3}{3} \right]_0^3 = 75 - \left(\frac{125}{3} - \frac{8}{3} \right) = 36$$

Figure 4.24



Alternatively,

$$\int_4^{25} (y^{1/2} - 2) dy = \left[\frac{2y^{3/2}}{3} - 2y \right]_4^{25} = \left(\frac{2 \cdot 125}{3} - 50 \right) - \left(\frac{16}{3} - 8 \right) = 36$$

(as above).

EXAMPLE

The quantity demanded and the corresponding price, under pure competition, are determined by the demand and supply functions $y = 16 - x^2$ and $y = 4 + x$, respectively. Determine the corresponding producers' surplus (see Figure 4.25).

$$y = 16 - x^2 = 4 + x$$

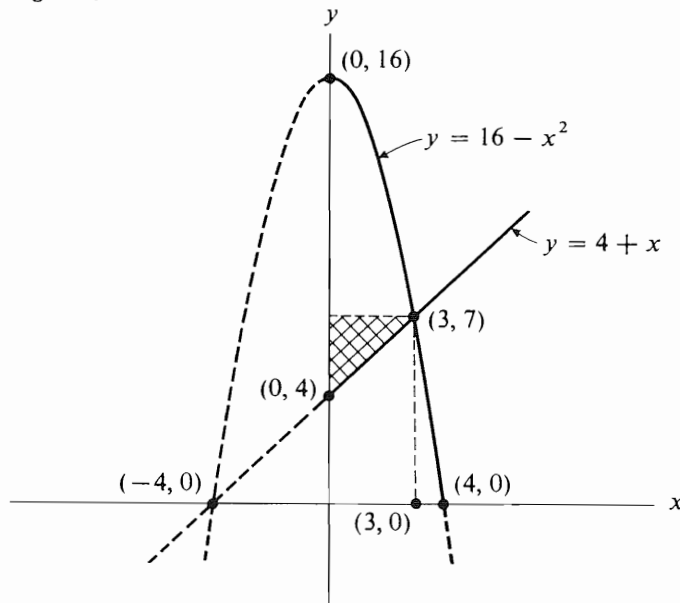
$$x^2 + x - 12 = 0$$

$$(x + 4)(x - 3) = 0$$

$$x = 3, -4$$

$$x_0 = 3, \quad y_0 = 7$$

Figure 4.25



$$\begin{aligned}
 \text{producers' surplus} &= (3)(7) - \int_0^3 (4 + x) \, dx \\
 &= 21 - \left[4x + \frac{x^2}{2} \right]_0^3 \\
 &= 21 - \left[(12 + \frac{9}{2}) - 0 \right] \\
 &= \frac{9}{2}
 \end{aligned}$$

NOTE: The area represented by $\int_0^3 (4 + x) \, dx$ could have been evaluated also by the formula for the area of a trapezoid, $A = \frac{1}{2}a(b_1 + b_2) = \frac{1}{2}(4 + 7) = \frac{33}{2}$.

EXAMPLE

The quantity demanded and the corresponding price, under pure competition, are determined by the demand and supply functions $y = 36 - x^2$ and $y = 6 + (x^2/4)$, respectively. Determine the corresponding consumers' surplus and producers' surplus (see Figure 4.26).

$$y = 36 - x^2 = 6 + \frac{x^2}{4}$$

$$5x^2 = 120$$

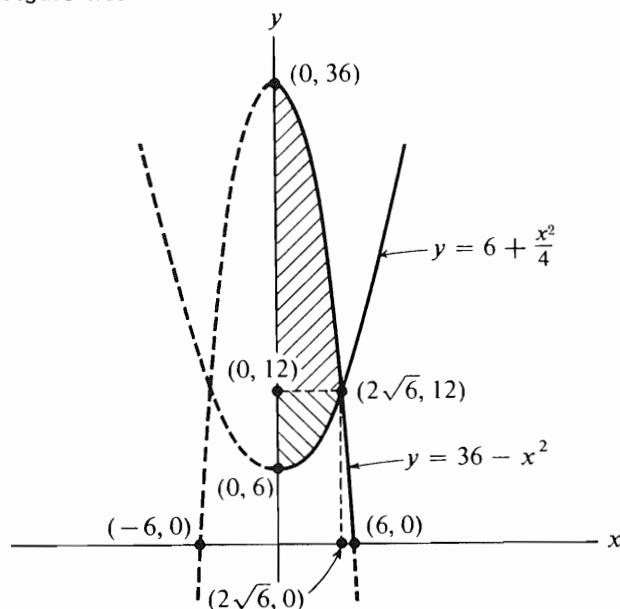
$$x^2 = 24$$

$$x = \pm 2\sqrt{6}$$

$$x_0 = 2\sqrt{6}, \quad y_0 = 12$$

$$\begin{aligned}
 \text{consumers' surplus} &= \int_0^{2\sqrt{6}} (36 - x^2) \, dx - (2\sqrt{6})(12) \\
 &= \left[36x - \frac{x^3}{3} \right]_0^{2\sqrt{6}} - 24\sqrt{6} \\
 &= (72\sqrt{6} - 16\sqrt{6}) - 0 - 24\sqrt{6} \\
 &= 32\sqrt{6} \approx 78.4
 \end{aligned}$$

Figure 4.26



$$\begin{aligned}
 \text{producers' surplus} &= (2\sqrt{6})(12) - \int_0^{2\sqrt{6}} \left(6 + \frac{x^2}{4}\right) dx \\
 &= 24\sqrt{6} - \left[6x + \frac{x^3}{12}\right]_0^{2\sqrt{6}} \\
 &= 24\sqrt{6} - (12\sqrt{6} + 4\sqrt{6}) - 0 \\
 &= 8\sqrt{6} \approx 19.6
 \end{aligned}$$

REVENUE VERSUS COST

Integration can be used to determine total profit or total net earnings in various contexts. In general, profit is maximized (assuming pure competition) when marginal revenue equals marginal cost. Total profit is the integral of marginal revenue minus marginal cost from zero quantity to the quantity for which profit is maximized.

EXAMPLE

Find the profit-maximizing output and the total profit at that point if the marginal revenue and marginal cost functions are given by

$$MR = 25 - 5x - 2x^2$$

$$MC = 15 - 2x - x^2$$

(see Figure 4.27).

If $MR - MC = 0$,

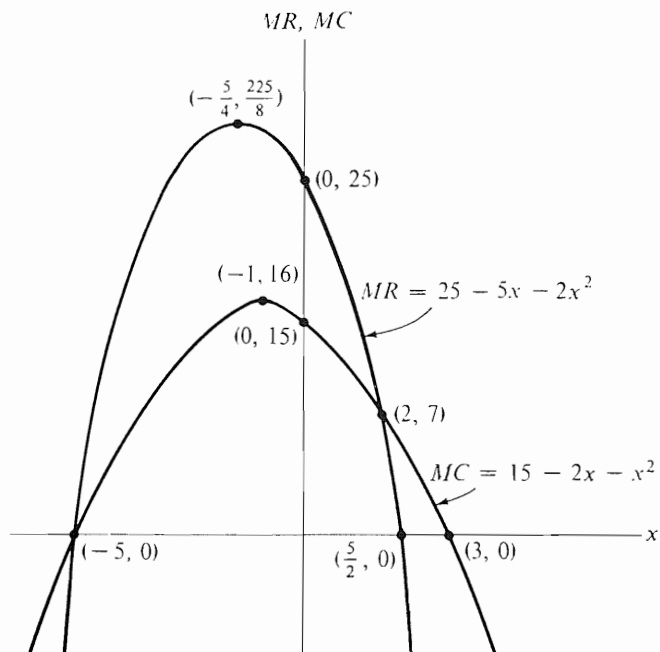
$$25 - 5x - 2x^2 - 15 + 2x + x^2 = 0$$

$$10 - 3x - x^2 = 0$$

$$(5 + x)(2 - x) = 0$$

$$x = -5, 2$$

Figure 4.27



The first derivative of $MR - MC$ is the second derivative of total profit and its sign thus indicates whether profit is maximized or minimized for a particular value of x .

$$\frac{d}{dx}(MR - MC) = \frac{d^2P}{dx^2} = -3 - 2x \quad \text{and} \quad \left. \frac{d^2P}{dx^2} \right|_{x=2} = -7$$

so profit is maximized for $x = 2$.

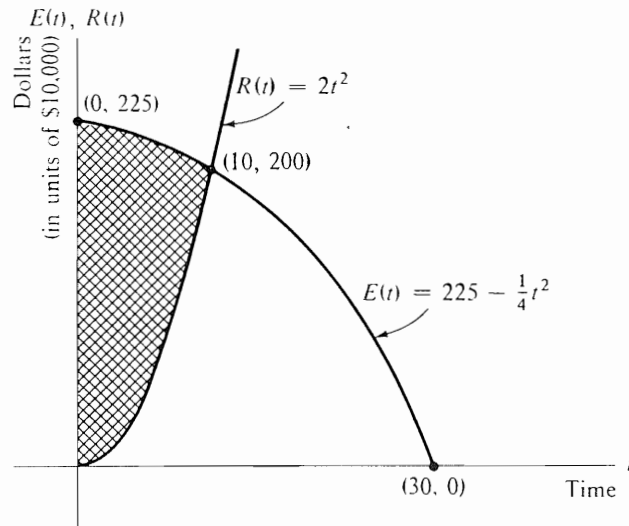
$$\begin{aligned} \text{total profit} &= \int_0^2 (10 - 3x - x^2) dx \\ &= \left[10x - \frac{3x^2}{2} - \frac{x^3}{3} \right]_0^2 \\ &= (20 - 6 - \frac{8}{3}) - 0 \\ &= \frac{34}{3} \end{aligned}$$

EXAMPLE

A manufacturing company has purchased a machine that has an output representing additional earnings (additional revenue minus additional cost of labor and materials) at time t of

$$E(t) = 225 - \frac{1}{4}t^2$$

Figure 4.28



where $E(t)$ is in units of \$10,000 and t is years. The additional repair and maintenance cost at time t is

$$R(t) = 2t^2$$

where $R(t)$ is in units of \$10,000 and t is years. First, suppose that the machine can be disposed of at any time with no cost or salvage value. Then the machine should be disposed of at the time when additional earnings equals additional cost of repair and maintenance (see Figure 4.28).

Additional earnings equals additional cost of repair and maintenance if

$$\begin{aligned} 225 - \frac{1}{4}t^2 &= 2t^2 \\ 225 &= \frac{9}{4}t^2 \\ t^2 &= 100 \\ t &= 10 \end{aligned}$$

Thus the machine should be disposed of after 10 years. The total net earnings (earnings minus cost of repair and maintenance) after 10 years is

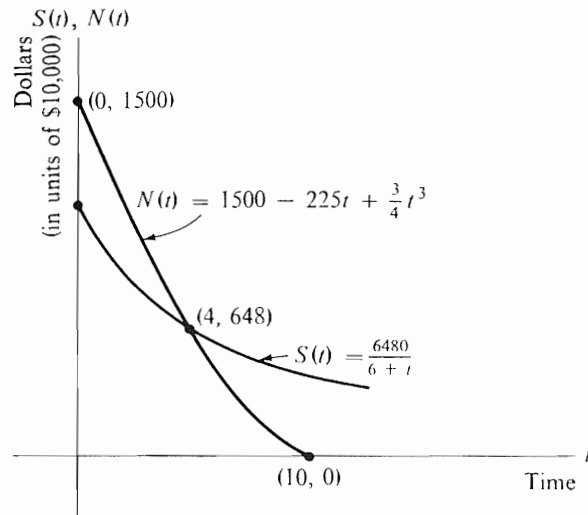
$$\begin{aligned} \int_0^{10} [E(t) - R(t)] dt &= \int_0^{10} (225 - \frac{9}{4}t^2) dt \\ &= [225t - \frac{3}{4}t^3]_0^{10} \\ &= (2250 - 750) - 0 \\ &= 1500 \quad \text{or } \$15,000,000 \end{aligned}$$

Now suppose the machine has a salvage value at time t of

$$S(t) = \frac{6480}{6+t}$$

where $S(t)$ is in units of \$10,000 and t is years. Then the company will maximize its net earnings if it disposes of the machine at time T when the net earnings after T equals the salvage value at T (see Figure 4.29).

Figure 4.29



Net earnings after T is

$$\begin{aligned} N(t) &= \int_T^{10} (225 - \frac{3}{4}t^2) dt \\ &= \left[225t - \frac{3t^3}{4} \right]_T^{10} \\ &= 1500 - 225T + \frac{3}{4}T^3 \end{aligned}$$

Net earnings after T equals salvage value at T if

$$\begin{aligned} \frac{6480}{6+T} &= 1500 - 225T + \frac{3}{4}T^3 \\ 6480 &= 9000 - 1350T + \frac{9}{2}T^3 + 1500T - 225T^2 + \frac{3}{4}T^4 \\ 0 &= 2520 + 150T - 225T^2 + \frac{9}{2}T^3 + \frac{3}{4}T^4 \\ &= (T-4)\left(\frac{3}{4}T^3 + \frac{1}{2}T^2 - 195T - 630\right) \end{aligned}$$

Thus the machine should be disposed of after 4 years.

EXAMPLE

A company is considering adding salesmen to its staff. The cost of employing additional salesmen is

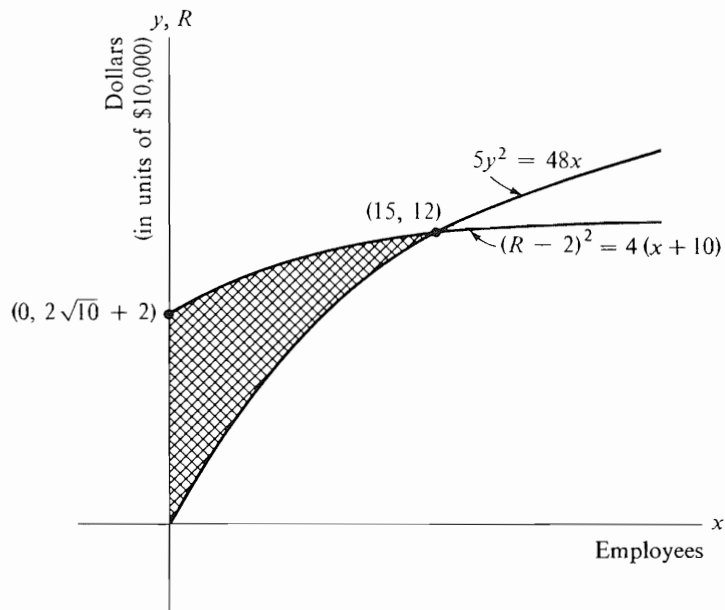
$$5y^2 = 48x$$

where y is cost in units of \$1,000 and x is number of additional salesmen employed, and the additional revenue is

$$(R-2)^2 = 4(x+10)$$

where R is revenue in units of \$1,000 and x is number of additional salesmen employed. (Assume the cost and revenue functions are continuous, although actually they are meaningful only for integer values of x .) The company should employ additional salesmen until the cost of doing so equals the additional revenue obtained (see Figure 4.30).

Figure 4.30



The cost of employing additional salesmen equals the additional revenue obtained if $R = y$. From the first equation,

$$x = \frac{5y^2}{48}$$

From the second equation,

$$x = \frac{(R-2)^2 - 40}{4}$$

Solving simultaneously and putting $R = y$,

$$\frac{5y^2}{48} = \frac{(y-2)^2 - 40}{4}$$

$$5y^2 = 12y^2 - 48y + 48 - 480$$

$$7y^2 - 48y - 432 = 0$$

$$(7y + 36)(y - 12) = 0$$

$$y = 12$$

$$x = 15$$

and 15 additional salesmen should be employed. The total resulting net revenue (total revenue minus cost) is

$$\begin{aligned} \int_0^{12} \left[\frac{5}{48}y^2 - \left(\frac{1}{4}(y-2)^2 - 10 \right) \right] dy &= \int_0^{12} \left[y + 9 - \frac{7}{48}y^2 \right] dy \\ &= \left[\frac{y^2}{2} + 9y - \frac{7}{144}y^3 \right]_0^{12} \\ &= (72 + 108 - 84) - 0 \\ &= 96 \quad \text{or } \$96,000 \end{aligned}$$

PROBLEMS

1. If the demand function is $y = 39 - x^2$, find the consumers' surplus if a. $x_0 = \frac{5}{2}$ and b. the commodity is free, that is, $y_0 = 0$.
2. If the demand function is $y = 16 - x^2$ and the supply function is $y = 2x + 1$, find consumers' surplus and producers' surplus under pure competition.
3. If the supply function is $y = \sqrt{9 + x}$ and $x_0 = 7$, find the producers' surplus.
4. If the supply function is $y = 4e^{x/3}$ and $x_0 = 3$, find the producers' surplus.
5. The demand and supply functions under pure competition are $y = \frac{1}{4}(9 - x)^2$ and $y = \frac{1}{4}(1 + 3x)$, respectively. If an additive tax of 3 per unit quantity is imposed on the commodity, determine the decrease in consumers' surplus.
6. Under monopoly, the quantity sold and the corresponding price are determined by the demand function $y = \frac{1}{4}(10 - x)^2$ and the total cost function $y = (x^3/4) + 5x$ so that profit is maximized. Determine the corresponding consumers' surplus.
7. Under monopoly, the quantity sold and the corresponding price are determined by the demand function $y = 20 - 4x^2$ and the marginal cost function $y' = 2x + 6$ so that profit is maximized. Determine the corresponding consumers' surplus.
8. If the demand function is that part of the equilateral hyperbola $y = [8/(x + 1)] - 2$ in the first quadrant, and the supply curve is $y = \frac{1}{2}(x + 3)$, find the consumers' surplus and the producers' surplus under pure competition.
9. Under monopoly, the quantity sold and the corresponding price are determined by the demand function $y = 45 - x^2$ and the marginal cost function $y' = 6 + (x^2/4)$ so that profit is maximized. Determine the corresponding consumers' surplus.
10. The demand and supply functions under pure competition are, respectively, $y = 14 - x^2$ and $y = 2x^2 + 2$; find a. consumers' surplus, and b. producers' surplus.
11. The demand function is $y = 20 - 3x^2$ and the supply function is $y = 2x^2$; find the consumers' surplus and the producers' surplus under pure competition.
12. The demand and supply functions under pure competition are, respectively, $y = 32 - 2x^2$ and $y = \frac{1}{3}x^2 + 2x + 5$; find a. consumers' surplus, and b. producers' surplus.
13. Find the profit-maximizing output and the corresponding total profit, assuming pure competition, if $MR = 20 - 2x$ and $MC = 4 + (x - 4)^2$.
14. The marginal revenue function is $MR = 25 - 3x$ and the marginal cost function is $MC = 25 - 7x + x^2$. Find the profit-maximizing output and the corresponding total profit under pure competition.
15. If $MR = 44 - 9x$ and $MC = 20 - 7x + 2x^2$, find the profit-maximizing output and the corresponding total profit under pure competition.
16. Assuming pure competition, find the profit-maximizing output and the corresponding total profit if $MR = 24 - 6x - x^2$ and $MC = 4 - 2x - x^2$.
17. If $MR = 15 - 5x$ and $MC = 10 - 3x + 3x^2$, find the profit-maximizing output and the corresponding total profit assuming pure competition.
18. A manufacturing company has purchased a machine whose output represents earnings at time t given by $y^2 = 6(t + 9)$, where y is in units of \$10,000 and t is years. The repair and maintenance cost at time t is given by $(y + 4)^2 = 8(t + 17)$, where y is in units of \$10,000 and t is years. Assuming the machine can be disposed of at any time with no cost or salvage value, how many years should it be kept to maximize total net earnings (earnings minus cost of repair and maintenance)?
19. A company is considering purchasing additional companion units for its computer. The savings (in time and mistakes) from added units is given by $y = 3x^2 + 11$, where y is in units of \$11,000 and x is number of units added. The cost of repair and maintenance is given by $y = 4x^2 + 2$, where y is in units of \$1000 and x is the number of units added. To maximize total net revenue (revenue minus cost of repair and maintenance), how many units should be added and what is the associated savings? Assume the units are added in a specified order and that the savings and cost curves are continuous.
20. A company is considering adding advertising personnel. The cost of adding such personnel is given by $y = \frac{1}{2}x$, where y is in units of \$5000 and x is the number of personnel added. Additional revenue from adding personnel is $R^2 = 4x$, where R is in units of \$5000 and x is number of personnel added. What number of advertising personnel should be added to

maximize profit (revenue minus cost) and what is the associated additional revenue? Assume continuous functions.

21. If net investment flow is $I(t) = 6t^{1/2}$, determine a. capital accumulation during the first year; b. capital accumulation during the first nine years; and c. total amount of capital after the first nine years if the initial capital is \$500.
22. If net investment flow is $I(t) = 4t^{1/3}$, determine a. capital accumulation during the first year; b. capital accumulation during the first eight years; and c. total amount of capital after the first eight years if the initial capital is \$500.

Answers to Odd-Numbered Problems

- | | |
|-----------------------|-------------------------------|
| 1. a. $31\frac{1}{4}$ | 13. $x = 6$; $P_{\max} = 36$ |
| b. $26\sqrt{13}$ | 15. $x = 3$; $P_{\max} = 45$ |
| 3. $\frac{10}{3}$ | 17. $x = 1$, $P_{\max} = 3$ |
| 5. $\frac{121}{12}$ | 19. 18 |
| 7. $\frac{8}{3}$ | 21. a. \$4.00 |
| 9. $16\sqrt{3}$ | b. \$108.00 |
| 11. c.s. = 16 | c. \$608.00 |
| p.s. = $\frac{32}{3}$ | |

4.6 Special Methods of Integration

In many cases, the integral of a given function cannot be obtained by reversing the rules of differentiation. A number of special methods are available for determining the integral in such cases. The special methods considered in this section are standard forms for integration, integration by parts, integration by partial fractions, integration by rationalizing substitution, and integration by miscellaneous substitution.

STANDARD FORMS FOR INTEGRATION

In differential calculus there is a general rule for differentiation, from which particular rules are derived. Unfortunately, there is no corresponding general rule for integration and the integral of any given expression must be obtained through knowledge of the results of differentiation. Thus integration is inherently more difficult than differentiation. In order to expedite the process of integrating a given expression, tables of known integrals, called *standard forms*, have been compiled. The simplest of these standard forms are obtained directly by reversing the corresponding rules for differentiation; additional standard forms are derived from the results of more complicated differentiation. Note that a formula for integration can be derived from the result of any differentiation; the most frequently occurring of such formulas are given in tables of standard forms.

In the following list of standard forms, the first 14 formulas are obtained by reversing the corresponding rules for differentiation (the first 5 of these are discussed at the beginning of the chapter); the last 12 formulas are examples of standard forms derived from the results of more complicated differentiation.

1. $\int dx = x + C.$
2. $\int K dx = K \int dx$, where K is any constant.
3. $\int [du + dv] = \int du + \int dv$, where $u = f(x)$ and $v = g(x)$ are differentiable functions of x .

$$4. \int x^n dx = \frac{x^{n+1}}{n+1} + C, n \neq -1.$$

$$5. \int u^n du = \frac{u^{n+1}}{n+1} + C, n \neq -1, \text{ where } u = f(x) \text{ is a differentiable function of } x.$$

$$6. \int \frac{1}{u} du = \ln u + C, \text{ where } u = f(x) \text{ is a differentiable function of } x.$$

$$7. \int e^u du = e^u + C, \text{ where } u = f(x) \text{ is a differentiable function of } x.$$

$$8. \int a^u du = \frac{a^u}{\ln a} + C, \text{ where } u = f(x) \text{ is a differentiable function of } x.$$

$$9. \int \sin u du = -\cos u + C.$$

$$10. \int \cos u du = \sin u + C.$$

$$11. \int \sec^2 u du = \tan u + C.$$

$$12. \int \csc^2 u du = -\cot u + C.$$

$$13. \int \sec u \tan u du = \sec u + C.$$

$$14. \int \csc u \cot u du = -\csc u + C.$$

$$15. \int \tan u du = -\ln \cos u + C = \ln \sec u + C.$$

$$16. \int \cot u du = \ln \sin u + C.$$

$$17. \int \sec u du = \ln (\sec u + \tan u) + C.$$

$$18. \int \csc u du = \ln (\csc u - \cot u) + C.$$

$$19. \int \frac{du}{a^2 - u^2} = \frac{1}{2a} \ln \frac{a+u}{a-u} + C, u^2 < a^2.$$

$$20. \int \frac{du}{u^2 - a^2} = \frac{1}{2a} \ln \frac{u-a}{u+a} + C, u^2 > a^2.$$

$$21. \int \frac{du}{\sqrt{u^2 + a^2}} = \ln (u + \sqrt{u^2 + a^2}) + C.$$

$$22. \int \sqrt{u^2 + a^2} du = \frac{u}{2} \sqrt{u^2 + a^2} + \frac{a^2}{2} \ln (u + \sqrt{u^2 + a^2}) + C.$$

$$23. \int ue^u du = e^u(u-1) + C.$$

$$24. \int \ln u \, du = u \ln u - u + C.$$

$$25. \int u^n \ln u \, du = u^{n+1} \left[\frac{\ln u}{n+1} - \frac{1}{(n+1)^2} \right] + C.$$

$$26. \int \frac{du}{u \ln u} = \ln (\ln u) + C.$$

The first step in the procedure for integrating a given expression consists of comparing it with various standard forms. If an expression is identical with a standard form, its integral is known; if an expression is not identical with a standard form, there are various methods by which it may in some cases be reduced to a standard form. Some of these methods are discussed in later sections.

EXAMPLES

Evaluate the following integrals.

$$\begin{aligned} \int \frac{x \, dx}{x^2 + 1} &= \frac{1}{2} \int \frac{2x \, dx}{x^2 + 1} \\ &= \frac{1}{2} \ln (x^2 + 1) + C \end{aligned} \quad \text{(standard form 6)}$$

$$\begin{aligned} \int x e^{-x^2} \, dx &= -\frac{1}{2} \int -2x e^{-x^2} \, dx \\ &= -\frac{1}{2} e^{-x^2} + C \end{aligned} \quad \text{(standard form 7)}$$

$$\begin{aligned} \int a^{2x-1} \, dx &= \frac{1}{2} \int 2a^{2x-1} \, dx \\ &= \frac{a^{2x-1}}{2 \ln a} + C \end{aligned} \quad \text{(standard form 8)}$$

$$\begin{aligned} \int \sin 3x \cos^2 3x \, dx &= -\frac{1}{3} \int (-3 \sin 3x)(\cos^2 3x) \, dx \\ &= -\frac{1}{3} \cos^3 3x + C \end{aligned} \quad \text{(standard form 5)}$$

$$\begin{aligned} \int \frac{\sin ax \, dx}{\sqrt{b + \cos ax}} &= -\frac{1}{a} \int (-a \sin ax)(b + \cos ax)^{-1/2} \, dx \\ &= -\frac{2}{a} \sqrt{b + \cos ax} + C \end{aligned} \quad \text{(standard form 5)}$$

$$\begin{aligned} \int \frac{\sec^2 by \, dy}{a + c \tan by} &= \frac{1}{bc} \int \frac{(bc \sec^2 by) \, dy}{a + c \tan by} \\ &= \frac{1}{bc} \ln (a + c \tan by) + C \end{aligned} \quad \text{(standard form 6)}$$

$$\int (x^2 - 1) \csc^2 (2x^3 - 6x) \, dx = -\frac{1}{6} \cot (2x^3 - 6x) + C \quad \text{(standard form 12)}$$

$$\begin{aligned} \int (e^x + 2x) \tan (e^x + x^2 + 5) \, dx &= -\ln \cos (e^x + x^2 + 5) + C \\ &= \ln \sec (e^x + x^2 + 5) + C \end{aligned} \quad \text{(standard form 15)}$$

$$\int \sec 4\pi x \tan 4\pi x \, dx = \frac{1}{4\pi} \sec 4\pi x + C \quad (\text{standard form 13})$$

$$\int (\theta + x) \csc^2 (\theta + x)^2 \, dx = -\frac{1}{2} \cot (\theta + x) + C \quad (\text{standard form 12})$$

$$\int \pi \csc \theta x \, dx = \frac{\pi}{\theta} \ln (\csc \theta x - \cot \theta x) + C \quad (\text{standard form 18})$$

$$\int 4 \csc 4\theta \cot 4\theta \, d\theta = -\csc \theta + C \quad (\text{standard form 14})$$

$$\int y^2 \sec (\theta + y^3) \, dy = \frac{1}{3} \ln [\sec (\theta + y^3) + \tan (\theta + y^3)] + C \quad (\text{standard form 17})$$

$$\int x \cot (x^2 + \theta x) \, dx = \frac{1}{2} \ln \sin (x^2 + \theta x) + C \quad (\text{standard form 16})$$

$$\int \frac{4 + 6x}{6 - (4x + 3x^2)^2} \, dx = \frac{1}{2\sqrt{6}} \ln \left(\frac{\sqrt{6} + 4x + 3x^2}{\sqrt{6} - 4x - 3x^2} \right) + C \quad (\text{standard form 20})$$

$$\int 4x^3 e^{x^2} \, dx = 2e^{x^2}(x^2 - 1) + C \quad (\text{standard form 23})$$

$$\int \frac{x^2}{x^3 \ln x^3} \, dx = \frac{1}{3} \ln (\ln x^3) + C \quad (\text{standard form 26})$$

$$\begin{aligned} \int (2x - 3)(x^2 - 3x - 1)^4 \ln (x^2 - 3x - 1) \, dx \\ = (x^2 - 3x - 1)^5 \left[\frac{\ln (x^2 - 3x - 1)}{5} - \frac{1}{25} \right] + C \end{aligned} \quad (\text{standard form 25})$$

$$\begin{aligned} \int (x^2 + 2x + 6)^{1/2} \, dx &= \int \sqrt{(x + 1)^2 + (\sqrt{5})^2} \, dx \\ &= \frac{x + 1}{2} (x^2 + 2x + 6)^{1/2} + \frac{5}{2} \ln [(x + 1) + (x^2 + 2x + 6)^{1/2}] + C \end{aligned} \quad (\text{standard form 22})$$

$$\begin{aligned} \int (x^2 + 2x) \ln (x^3 + 3x^2 + 14) \, dx \\ = \frac{1}{3} [(x^3 + 3x^2 + 14) \ln (x^3 + 3x^2 + 14) - (x^3 + 3x^2 + 14)] + C \end{aligned} \quad (\text{standard form 24})$$

PROBLEMS

Evaluate the following integrals.

1. $\int x e^{x^2} \, dx$

4. $\int \sec^2 \frac{x}{a} \tan \frac{x}{a} \, dx$

2. $\int \sin^2 x \cos x \, dx$

5. $\int \left(\frac{\sec 2x}{1 + \tan 2x} \right)^2 \, dx$

3. $\int \sin ax \cos ax \, dx$

6. $\int (x^2 + 1)5^{x^3 + 3x} \, dx$

7. $\int \frac{\cos x \, dx}{1 + \sin x}$

14. $\int \frac{\sin 2x}{\cos^2 2x} \, dx$

8. $\int e^{\cos z} \sin z \, dz$

15. $\int (\sin x + \cos x) \, dx$

9. $\int \frac{dx}{\sin^2 ax}$

16. $\int (3x^2 + 5 \cos x) \, dx$

10. $\int \frac{du}{\cos^2 u}$

17. $\int (3x + e^x) \, dx$

11. $\int \sec ax \, dx$

18. $\int (e^x - e^{-x}) \, dx$

12. $\int \frac{dt}{1 + \cos t}$

19. $\int 2x \sec x^2 \tan x^2 \, dx$

13. $\int 2x \cos x^2 \, dx$

20. $\int (x^3 + 3x)(3x^2 + 3)e^{x^3+3x} \, dx$

Answers to Odd-Numbered Problems

1. $\frac{1}{2}e^{x^2} + C$

11. $\frac{1}{a} \ln (\sec ax + \tan ax) + C$

3. $\frac{\sin^2 ax}{2a} + C$

13. $-\sin x^2 + C$

5. $-\frac{1}{2(1 + \tan 2x)} + C$

15. $\sin x - \cos x + C$

7. $\ln (1 + \sin x) + C$

17. $\frac{3}{2}x^2 + e^x + C$

9. $-\frac{1}{a} \cot ax + C$

19. $\sec x^2 + C$

INTEGRATION BY PARTS

When an expression involving products or logarithms cannot be evaluated directly using standard forms, one of the most useful techniques for transforming it to a standard form is the formula for integration by parts. This formula is based on the inverse of the formula for differentiation of a product.

If u and v are both functions of x , then from the formula for the differentiation of a product,

$$\frac{d}{dx}(uv) = u \frac{dv}{dx} + v \frac{du}{dx}$$

or

$$u \frac{dv}{dx} = \frac{d}{dx}(uv) - v \frac{du}{dx}$$

and, integrating with respect to x ,

$$\int u \, dv = uv - \int v \, du$$

This is the *formula for integration by parts*. The usefulness of this formula depends on the appropriate choice of u and dv , so that $\int v \, du$ and $\int dv$ can be evaluated, even though $\int u \, dv$ cannot.

Unfortunately, there is no general rule for separating a given expression into two factors u and dv so that the formula for integration by parts can be applied. However, note that

1. dx is always a part of dv .
2. dv must be integrable.
3. When the expression to be integrated is the product of two functions, it is usually advisable to choose the most complicated looking one that can be integrated as part of dv in order to make $\int v du$ as easily integrable as possible.

It may be necessary to apply the formula for integration by parts more than once, as in example (e) below.

EXAMPLES

Evaluate the following integrals.

$$\begin{aligned}
 \text{(a)} \quad \int x e^{ax} dx &= \frac{x e^{ax}}{a} - \int \frac{e^{ax}}{a} dx + C \\
 &= \frac{x e^{ax}}{a} - \frac{e^{ax}}{a^2} + C & u = x & \quad dv = e^{ax} dx \\
 &= \frac{e^{ax}}{a} \left(x - \frac{1}{a} \right) + C & du = dx & \quad v = \frac{e^{ax}}{a}
 \end{aligned}$$

(Note that, alternatively, standard form 23 could be used.)

$$\begin{aligned}
 \text{(b)} \quad \int \ln x dx &= x \ln x - \int x \left(\frac{1}{x} \right) dx + C \\
 &= x \ln x - x + C & u = \ln x & \quad dv = dx \\
 &= x(\ln x - 1) + C & du = \frac{1}{x} dx & \quad v = x
 \end{aligned}$$

(Note that, alternatively, standard form 24 could be used.)

$$\begin{aligned}
 \text{(c)} \quad \int x \sin \frac{x}{a} dx &= -ax \cos \frac{x}{a} + \int a \cos \frac{x}{a} dx + C \\
 &= -ax \cos \frac{x}{a} + a^2 \sin \frac{x}{a} + C & u = x & \quad dv = \sin \frac{x}{a} dx \\
 &= a^2 \sin \frac{x}{a} - ax \cos \frac{x}{a} + C & du = dx & \quad v = -a \cos \frac{x}{a}
 \end{aligned}$$

$$\begin{aligned}
 \text{(d)} \quad \int e^x \cos x dx &= e^x \sin x - \int e^x \sin x dx + C \\
 &= e^x \sin x - \left[-e^x \cos x + \int e^x \cos x dx \right] + C \\
 &= e^x \sin x + e^x \cos x - \int e^x \cos x dx + C \\
 \int e^x \cos x dx &= \frac{1}{2} e^x (\sin x + \cos x) + C
 \end{aligned}$$

$$u = e^x \quad dv = \cos x dx$$

$$du = e^x \quad v = \sin x$$

$$u = e^x \quad dv = \sin x dx$$

$$du = e^x \quad v = -\cos x$$

$$\begin{aligned}
 \text{(e)} \quad \int x^2 e^{-x} dx &= -x^2 e^{-x} + 2 \int x e^{-x} dx + C \\
 &= -x^2 e^{-x} + 2 \left[-x e^{-x} + \int e^{-x} dx \right] + C \\
 &= -x^2 e^{-x} - 2x e^{-x} - 2e^{-x} + C \\
 &= -e^{-x}(x^2 + 2x + 2) + C
 \end{aligned}$$

$$u = x^2 \quad dv = e^{-x} dx$$

$$du = 2x dx \quad v = -e^{-x}$$

$$u = x \quad dv = e^{-x} dx$$

$$du = dx \quad v = e^{-x}$$

$$\begin{aligned}
 \text{(f)} \quad \int \frac{\ln(x+1) dx}{\sqrt{x+1}} &= 2(x+1)^{1/2} \ln(x+1) - 2 \int \frac{1}{x+1} (x+1)^{1/2} dx + C \\
 &= 2(x+1)^{1/2} \ln(x+1) - 4(x+1)^{1/2} + C \\
 &= 2(x+1)^{1/2} [\ln(x+1) - 2] + C
 \end{aligned}$$

$$u = \ln(x+1) \quad dv = (x+1)^{-1/2} dx$$

$$du = \frac{dx}{x+1} \quad v = 2(x+1)^{1/2}$$

(Note that, alternatively, standard form 25 could be used.)

As suggested by examples (a), (b), and (f) above, the formula for integration by parts is the basis of a number of standard forms for integration.

PROBLEMS

Evaluate the following integrals.

$$1. \int \frac{x e^x dx}{(1+x)^2}$$

$$2. \int x e^{-x} dx$$

$$3. \int x^2 e^x dx$$

$$4. \int x e^{2x} dx$$

$$5. \int t \ln t dt$$

$$6. \int z^2 e^{-3z} dz$$

$$7. \int e^x (x+1)^2 dx$$

$$8. \int (x^{1/2} + x^{1/4})^2 dx$$

$$9. \int \frac{(x+1) dx}{(x+1)^2 + (a+1)^2}$$

$$10. \int (x+1) \ln x dx$$

$$11. \int (x^2 + 3x + 4)^3 (2x + 3) dx$$

$$12. \int e^{3x^2+6x+10} (x+1) dx$$

$$13. \int (x^{1/3} + x^{2/3})^2 dx$$

$$14. \int \frac{6x^2 + 8x + 8}{x^3 + 2x^2 + 4x} dx$$

$$15. \int \frac{(x^3 + x) dx}{\sqrt[4]{x^4 + 2x^2 + 1}}$$

$$16. \int \left(x^2 + \frac{1}{x} + 6x \right) dx$$

$$17. \int x e^{-3x} dx$$

$$18. \int \frac{x^3 + 2}{x^4 + 8x + 10} dx$$

19. $\int \frac{\sin v \, dv}{1 + \cos v}$

20. $\int \frac{dx}{1 + \sin x}$

21. $\int u \sin u^2 \, du$

22. $\int (x + \sin 2x) \, dx$

23. $\int xa^x \, dx$

24. $\int x^n \ln x \, dx$

25. $\int \frac{\sin x \, dx}{\sqrt{2 - \cos x}}$

26. $\int \frac{\sec^2 \theta \, d\theta}{\sqrt{1 + 2 \tan \theta}}$

27. $\int x \cos x \, dx$

28. $\int e^{-ax} \sin nx \, dx$

29. $\int \theta \sec^2 \theta \, d\theta$

30. $\int y^2 \sin ny \, dy$

Answers to Odd-Numbered Problems

1. $\frac{e^x}{1+x} + C$

3. $e^x(x^2 - 2x + 2) + C$

5. $\frac{1}{2}t^2 \ln t - \frac{1}{4}t^2 + C$

7. $e^x(x^2 + 1) + C$

9. $\frac{1}{2} \ln [(x+1)^2 + (a+1)^2] + C$

11. $\frac{1}{4}(x^2 + 3x + 4)^4 + C$

13. $\frac{3}{5}x^{5/3} + x^2 + \frac{3}{7}x^{7/3} + C$

15. $\frac{1}{3}(x^4 + 2x^2 + 1)^{3/4} + C$

17. $-\frac{1}{9}e^{-3x}(3x+1) + C$

19. $-\ln(1 + \cos v) + C$

21. $-\frac{1}{2} \cos u^2 + C$

23. $\frac{a^x}{\ln a} \left(x - \frac{1}{\ln a} \right) + C$

25. $2\sqrt{2 - \cos x} + C$

27. $x \sin x + \cos x + C$

29. $\theta \tan \theta + \ln \cos \theta + C$

INTEGRATION BY PARTIAL FRACTIONS

A rational algebraic function, by definition, can be expressed as the quotient of two polynomials. In theory, every rational function has an integral that can be expressed in terms of elementary functions. If a rational function cannot be integrated directly, the method of partial fractions can frequently be used to transform the rational fraction into a sum of simpler functions which can then be integrated by standard forms. The method of partial fractions is appropriate *only* for proper fractions, that is, for fractions in which the polynomial in the numerator is of lower degree than the polynomial in the denominator. Any improper fraction, that is, a fraction in which the polynomial in the numerator is of the same or higher degree than the polynomial in the denominator, can be reduced by division to the sum of a polynomial (easily integrated) and a proper fraction (integrated by the method of partial fractions). The method of integration by partial fractions consists of the following steps:

1. Express the denominator of the fraction as the product of linear factors of the form $ax + b$ and irreducible quadratic factors of the form $ax^2 + bx + c$; although this is not always easy in practice, it is always possible in theory for any polynomial in x with real coefficients.

2. Determine the form of the partial fractions; several cases arise, depending on the nature of the factors occurring in the denominator:

*Factor Occurring
in Denominator*

Corresponding Partial Fraction

- (a) Distinct linear
factor: $ax + b$

$$\frac{A}{ax + b}$$

where A is a constant to be determined

- (b) Repeated linear
factor: $(ax + b)^n$

$$\frac{A_1}{ax + b} + \frac{A_2}{(ax + b)^2} + \cdots + \frac{A_n}{(ax + b)^n}$$

where A_1, \dots, A_n are constants to be determined

- (c) Distinct quad-
ratic factor:
 $ax^2 + bx + c$

$$\frac{Ax + B}{ax^2 + bx + c}$$

where A and B are constants to be determined

- (d) Repeated quad-
ratic factor:
 $(ax^2 + bx + c)^n$

$$\frac{A_1x + B_1}{ax^2 + bx + c} + \frac{A_2x + B_2}{(ax^2 + bx + c)^2} + \cdots + \frac{A_nx + B_n}{(ax^2 + bx + c)^n}$$

where A_i and B_i , $i = 1, 2, \dots, n$
are constants to be determined

3. Determine the constants occurring in the numerators of the partial fractions. When a rational fraction is separated into partial fractions, the resulting equation is an identity, that is, it is true for all meaningful values of the variables. The method for evaluating the constants occurring in partial fractions, as illustrated in the example below, is based on the algebraic theorem which states that if two polynomials of the same degree are identical, the coefficients of like powers of the variable in both polynomials must be equal.
4. Integrate the partial fractions using standard forms.

EXAMPLES

Evaluate the following integrals.

$$(a) \int \frac{(x+3) dx}{x^2 + 3x + 2} = \int \frac{(x+3) dx}{(x+1)(x+2)} = \int \left[\frac{A}{x+1} + \frac{B}{x+2} \right] dx$$

$$x + 3 = A(x+2) + B(x+1)$$

$$= (A+B)x + (2A+B)$$

so, equating coefficients,

$$A + B = 1$$

$$2A + B = 3$$

$$A = 2$$

$$B = -1$$

and

$$\begin{aligned}\int \frac{(x+3) dx}{(x+1)(x+2)} &= \int \frac{2 dx}{x+1} - \int \frac{dx}{x+2} \\ &= 2 \ln(x+1) - \ln(x+2) + C \\ &= \ln \left[\frac{(x+1)^2}{x+2} \right] + C\end{aligned}$$

$$\begin{aligned}\text{(b)} \int \frac{(x^2 - 3x - 8) dx}{x^2 - 2x + 1} &= \int \left[1 - \frac{x+9}{(x-1)^2} \right] dx = x - \int \left[\frac{A}{x-1} + \frac{B}{(x-1)^2} \right] dx \\ x+9 &= A(x-1) + B \\ &= Ax - (A-B)\end{aligned}$$

so, equating coefficients,

$$A = 1$$

$$A - B = -9$$

$$B = 10$$

and

$$\begin{aligned}\int \frac{(x^2 - 3x - 8) dx}{x^2 - 2x + 1} &= x - \int \frac{dx}{x-1} - \int \frac{10 dx}{(x-1)^2} \\ &= x - \ln(x-1) + \frac{10}{x-1} + C\end{aligned}$$

$$\begin{aligned}\text{(c)} \int \frac{t dt}{t^4 + 6t^2 + 5} &= \int \frac{t dt}{(t^2 + 5)(t^2 + 1)} = \int \left[\frac{A_1 t + B_1}{t^2 + 5} + \frac{A_2 t + B_2}{t^2 + 1} \right] dt \\ t &= (A_1 t + B_1)(t^2 + 1) + (A_2 t + B_2)(t^2 + 5) \\ &= (A_1 + A_2)t^3 + (B_1 + B_2)t^2 + (A_1 + 5A_2)t + (B_1 + 5B_2)\end{aligned}$$

so

$$A_1 + A_2 = 0$$

$$B_1 + B_2 = 0$$

$$A_1 + 5A_2 = 1$$

$$B_1 + 5B_2 = 0$$

From the first and third equations,

$$A_2 = \frac{1}{4}$$

$$A_1 = -\frac{1}{4}$$

From the second and fourth equations,

$$B_1 = B_2 = 0$$

and

$$\begin{aligned}\int \frac{t dt}{(t^2 + 5)(t^2 + 1)} &= -\frac{1}{4} \int \frac{t dt}{t^2 + 5} + \frac{1}{4} \int \frac{t dt}{t^2 + 1} \\ &= -\frac{1}{8} \ln(t^2 + 5) + \frac{1}{8} \ln(t^2 + 1) + C \\ &= \frac{1}{8} \ln \left(\frac{t^2 + 1}{t^2 + 5} \right) + C\end{aligned}$$

$$(d) \int \frac{dz}{z(z^2 + 1)^2} = \int \left[\frac{A_0}{z} + \frac{A_1 z + B_1}{z^2 + 1} + \frac{A_2 z + B_2}{(z^2 + 1)^2} \right] dz$$

$$\begin{aligned} 1 &= A_0(z^2 + 1)^2 + (A_1 z + B_1)z(z^2 + 1) + (A_2 z + B_2)z \\ &= A_0(z^4 + 2z^2 + 1) + A_1(z^4 + z^2) + B_1(z^3 + z) + A_2 z^2 + B_2 z \\ &= (A_0 + A_1)z^4 + B_1 z^3 + (2A_0 + A_1 + A_2)z^2 + (B_1 + B_2)z + A_0 \end{aligned}$$

so, equating coefficients,

$$A_0 + A_1 = 0$$

$$B_1 = 0$$

$$2A_0 + A_1 + A_2 = 0$$

$$B_1 + B_2 = 0$$

$$A_0 = 1$$

Substituting $A_0 = 1$ and solving the other equations,

$$A_1 = -1$$

$$A_2 = -1$$

$$B_2 = 0$$

and

$$\begin{aligned} \int \frac{dz}{z(z^2 + 1)^2} &= \int \frac{dz}{z} - \int \frac{z dz}{z^2 + 1} - \int \frac{z dz}{(z^2 + 1)^2} \\ &= \ln z - \frac{1}{2} \ln(z^2 + 1) + \frac{1}{2(z^2 + 1)} + C \\ &= \frac{1}{2} \ln \left(\frac{z^2}{z^2 + 1} \right) + \frac{1}{2(z^2 + 1)} + C \end{aligned}$$

$$(e) \int \frac{dx}{x^3 + 5x^2 + 4x} = \int \frac{dx}{x(x+1)(x+4)} = \int \left[\frac{A}{x} + \frac{B}{x+1} + \frac{C}{x+4} \right] dx$$

$$\begin{aligned} 1 &= A(x+1)(x+4) + Bx(x+4) + Cx(x+1) \\ &= A(x^2 + 5x + 4) + B(x^2 + 4x) + C(x^2 + x) \\ &= (A + B + C)x^2 + (5A + 4B + C)x + 4A \end{aligned}$$

so, equating coefficients,

$$A + B + C = 0$$

$$5A + 4B + C = 0$$

$$4A = 1$$

$$A = \frac{1}{4}$$

Substituting $A = \frac{1}{4}$ and solving the first two equations,

$$4A + 3B = 0$$

$$B = -\frac{1}{3}$$

$$C = \frac{1}{12}$$

and

$$\begin{aligned}\int \frac{dx}{x(x+1)(x+4)} &= \frac{1}{4} \int \frac{dx}{x} - \frac{1}{3} \int \frac{dx}{x+1} + \frac{1}{12} \int \frac{dx}{x+4} \\ &= \frac{1}{4} \ln x - \frac{1}{3} \ln(x+1) + \frac{1}{12} \ln(x+4) + C \\ &= \frac{1}{12} \ln \left[\frac{x^3(x+4)}{(x+1)^4} \right] + C\end{aligned}$$

$$\begin{aligned}\text{(f)} \int \frac{(x^3 - 2x) dx}{x^4 - 81} &= \int \frac{(x^3 - 2x) dx}{(x-3)(x+3)(x^2+9)} \\ &= \int \left[\frac{A}{x-3} + \frac{B}{x+3} + \frac{Cx+D}{x^2+9} \right] dx \\ x^3 - 2x &= A(x+3)(x^2+9) + B(x-3)(x^2+9) + (Cx+D)(x^2-9) \\ &= A(x^3 + 3x^2 + 9x + 27) + B(x^3 - 3x^2 + 9x - 27) \\ &\quad + C(x^3 - 9x) + D(x^2 - 9) \\ &= (A+B+C)x^3 + (3A-3B+D)x^2 + (9A+9B-9C)x \\ &\quad + (27A-27B-9D)\end{aligned}$$

so, equating coefficients,

$$A + B + C = 1$$

$$3A - 3B + D = 0$$

$$9A + 9B - 9C = -2$$

$$27A - 27B - 9D = 0$$

From the first and third equations,

$$18C = 11$$

$$C = \frac{11}{18}$$

From the second and fourth equations,

$$18D = 0$$

$$D = 0$$

Substituting $C = \frac{11}{18}$, $D = 0$, and solving the first two equations,

$$A + B = \frac{7}{18}$$

$$3A - 3B = 0$$

$$6A = \frac{7}{18}$$

$$A = \frac{7}{36}$$

$$B = \frac{7}{36}$$

and

$$\begin{aligned}\int \frac{(x^3 - 2x) dx}{(x-3)(x+3)(x^2+9)} &= \frac{7}{36} \int \frac{dx}{x-3} + \frac{7}{36} \int \frac{dx}{x+3} + \frac{11}{18} \int \frac{x dx}{x^2+9} \\ &= \frac{7}{36} \ln(x-3) + \frac{7}{36} \ln(x+3) + \frac{11}{36} \ln(x^2+9) + C \\ &= \frac{1}{36} \ln[(x-3)^7(x+3)^7(x^2+9)^{11}] + C \\ &= \frac{1}{36} \ln[(x^2-9)^7(x^2+9)^{11}] + C\end{aligned}$$

NOTE: In this particular problem the correct solution is also obtained using the factors $(x^2 - 9)(x^2 + 9)$; however, in general, *irreducible* quadratic factors and linear factors must be used.

PROBLEMS

Evaluate each of the following integrals using partial fractions.

$$1. \int \frac{(4x - 2) dx}{x^3 - x^2 - 2x}$$

$$5. \int \frac{(4x^2 + 6) dx}{x^3 + 3x}$$

$$2. \int \frac{(4x^3 + 2x^2 + 1) dx}{4x^3 - x}$$

$$6. \int \frac{(x^3 + 3x) dx}{(x^2 + 1)^2}$$

$$3. \int \frac{z^2 dz}{(z - 1)^3}$$

$$7. \int \frac{t^5 dt}{(t^2 + 4)^2}$$

$$4. \int \frac{(y^4 - 8) dy}{y^3 + 2y^2}$$

$$8. \int \frac{(2t^2 - 8t - 8) dt}{(t - 2)(t^2 + 4)}$$

Answers to Odd-Numbered Problems

$$1. \ln \left[\frac{x(x - 2)}{(x + 1)^2} \right] + C$$

$$5. \ln [x^2(x^2 + 3)] + C$$

$$3. \ln(z - 1) + \frac{3 - 4z}{2(z - 1)^2} + C$$

$$7. \frac{t^2}{2} - 4 \ln(t^2 + 4) - \frac{8}{t^2 + 4} + C$$

INTEGRATION BY A RATIONALIZING SUBSTITUTION

As discussed in the preceding section, all rational functions are integrable in terms of elementary functions; however, only a relatively small number of algebraic functions which are not rational can be integrated in terms of elementary functions. In some cases, by substitution of a new variable, functions which are not rational can be transformed into equivalent functions which are rational. These equivalent functions are integrable either by partial fractions or directly by standard forms. The method of integrating a function that is not rational by transforming it into a rational function by substitution is called *integration by rationalization*. This method is useful for various classes of functions; two of the most common of these are discussed in the following sections.

1. An expression involving fractional powers of x only can be transformed into a rational form by the substitution

$$x = z^n$$

where n is the least common denominator of the fractional exponents of x .

2. An expression involving fractional powers of $(a + bx)$ only can be transformed into a rational form by the substitution

$$a + bx = z^n$$

where n is the least common denominator of the fractional exponents of $(a + bx)$.

EXAMPLES

Evaluate the following integrals.

$$(a) \int \frac{x^{1/2} dx}{1 + x^{3/4}}$$

involves fractional powers of x whose least common denominator is 4. Substituting,

$$x = z^4$$

$$dx = 4z^3 dz$$

$$\begin{aligned} \int \frac{x^{1/2} dx}{1 + x^{3/4}} &= \int \frac{z^2 \cdot 4z^3 dz}{1 + z^3} \\ &= 4 \int \frac{z^5 dz}{z^3 + 1} \\ &= 4 \int \left[z^2 - \frac{z^2}{z^3 + 1} \right] dz \\ &= \frac{4}{3} z^3 - \frac{4}{3} \ln(z^3 + 1) + C \end{aligned}$$

Substituting to obtain a function of x ,

$$\int \frac{x^{1/2} dx}{1 + x^{3/4}} = \frac{4}{3} x^{3/4} - \frac{4}{3} \ln(x^{3/4} + 1) + C$$

$$(b) \int \frac{x dx}{(a + bx)^{3/2}}$$

involves a fractional power of $a + bx$ whose least common denominator is 2.

Substituting,

$$a + bx = z^2$$

$$b dx = 2z dz$$

$$dx = \frac{2}{b} z dz$$

$$\begin{aligned} \int \frac{x dx}{(a + bx)^{3/2}} &= \frac{2}{b^2} \int \frac{(z^2 - a) \cdot z dz}{z^3} \\ &= \frac{2}{b^2} \int \frac{(z^3 - az) dz}{z^3} \\ &= \frac{2}{b^2} \int \left[1 - \frac{a}{z^2} \right] dz \\ &= \frac{2}{b^2} \left[z + \frac{a}{z} \right] + C \\ &= \frac{2}{b^2} \left[\frac{z^2 + a}{z} \right] + C \end{aligned}$$

Substituting to obtain a function of x ,

$$\int \frac{x dx}{(a + bx)^{3/2}} = \frac{2}{b^2} \left[\frac{2a + bx}{(a + bx)^{1/2}} \right] + C$$

$$(c) \int \frac{(5x + 9) dx}{(x - 9)x^{3/2}}$$

involves a fractional power of x whose least common denominator is 2. Substituting,

$$x = z^2$$

$$dx = 2z dz$$

$$\begin{aligned} \int \frac{(5x + 9) dx}{(x - 9)x^{3/2}} &= \int \frac{(5z^2 + 9) \cdot 2z dz}{(z^2 - 9)z^3} \\ &= 2 \int \frac{(5z^2 + 9) dz}{z^2(z - 3)(z + 3)} \end{aligned}$$

Integrating using partial fractions,

$$\begin{aligned}\int \frac{(5x+9) dx}{(x-9)x^{3/2}} &= 2 \int \left[\frac{A}{z} + \frac{B}{z^2} + \frac{C}{z-3} + \frac{D}{z+3} \right] dz \\ 5z^2 + 9 &= Az(z^2-9) + B(z^2-9) + Cz^2(z+3) + Dz^2(z-3) \\ &= A(z^3-9z) + B(z^2-9) + C(z^3+3z^2) + D(z^3-3z^2)\end{aligned}$$

so

$$\begin{aligned}A + C + D &= 0 \\ B + 3C - 3D &= 5 \\ -9A &= 0 \\ -9B &= 9 \\ A &= 0 \\ B &= -1\end{aligned}$$

Substituting $A = 0$, $B = -1$, and solving the first two equations,

$$\begin{aligned}C + D &= 0 \\ 3C - 3D &= 6 \\ 6C &= 6 \\ C &= 1 \\ D &= -1\end{aligned}$$

and

$$\begin{aligned}\int \frac{(5x+9) dx}{(x-9)x^{3/2}} &= 2 \left[-\int \frac{dz}{z^2} + \int \frac{dz}{z-3} - \int \frac{dz}{z+3} \right] \\ &= 2 \left[\frac{1}{z} + \ln(z-3) - \ln(z+3) \right] + C \\ &= 2 \left[\frac{1}{z} + \ln \left(\frac{z-3}{z+3} \right) \right] + C\end{aligned}$$

Substituting to obtain a function of x ,

$$\int \frac{(5x+9) dx}{(x-9)x^{3/2}} = 2 \left[\frac{1}{x^{1/2}} + \ln \left(\frac{x^{1/2}-3}{x^{1/2}+3} \right) \right] + C$$

$$(d) \int \frac{dx}{x-x^{4/3}}$$

involves a fractional power of x whose least common denominator is 3. Substituting,

$$\begin{aligned}x &= z^3 \\ dx &= 3z^2 dz \\ \int \frac{dx}{x-x^{4/3}} &= \int \frac{3z^2 dz}{z^3-z^4} \\ &= 3 \int \frac{dz}{z(1-z)}\end{aligned}$$

Integrating using partial fractions,

$$\begin{aligned}\int \frac{dx}{x-x^{4/3}} &= 3 \int \left[\frac{A}{z} + \frac{B}{1-z} \right] dz \\ 1 &= A(1-z) + Bz\end{aligned}$$

so

$$-A + B = 0$$

$$A = 1$$

$$B = 1$$

and

$$\begin{aligned}\int \frac{dx}{x - x^{4/3}} &= 3 \left[\int \frac{dz}{z} + \int \frac{dz}{1-z} \right] \\ &= 3[\ln z - \ln(1-z)] + C \\ &= 3 \ln \left(\frac{z}{1-z} \right) + C\end{aligned}$$

Substituting to obtain a function of x ,

$$\int \frac{dx}{x - x^{4/3}} = 3 \ln \left(\frac{x^{1/3}}{1 - x^{1/3}} \right) + C$$

$$(e) \int y(a+y)^{1/3} dy$$

involves a fractional power of $(a+y)$ whose least common denominator is 3.
Substituting,

$$a+y = z^3$$

$$dy = 3z^2 dz$$

$$\begin{aligned}\int y(a+y)^{1/3} dy &= \int (z^3 - a) \cdot z \cdot 3z^2 dz \\ &= 3 \int (z^6 - az^3) dz \\ &= 3 \left[\frac{z^7}{7} - \frac{az^4}{4} \right] + C \\ &= \frac{3}{28} [z^4(4z^3 - 7a)] + C\end{aligned}$$

Substituting to obtain a function of x ,

$$\begin{aligned}\int y(a+y)^{1/3} dy &= \frac{3}{28} [(a+y)^{4/3} (4a+4y-7a)] + C \\ &= \frac{3}{28} [(a+y)^{4/3} (4y-3a)] + C\end{aligned}$$

$$(f) \int \frac{x^2 dx}{(4x+1)^{5/2}}$$

involves a fractional power of $(4x+1)$ whose least common denominator is 2.
Substituting,

$$4x+1 = z^2$$

$$4 dx = 2z dz$$

$$dx = \frac{1}{2} z dz$$

$$\begin{aligned}
\int \frac{x^2 dx}{(4x+1)^{5/2}} &= \int \frac{\frac{1}{16}(z^2-1)^2 \cdot \frac{1}{2}z dz}{z^5} \\
&= \frac{1}{32} \int \frac{(z^4 - 2z^2 + 1) dz}{z^4} \\
&= \frac{1}{32} \int \left[1 - \frac{2}{z^2} + \frac{1}{z^4} \right] dz \\
&= \frac{1}{32} \left[z + \frac{2}{z} - \frac{1}{3z^3} \right] + C \\
&= \frac{1}{32} \left[\frac{3z^4 + 6z^2 - 1}{3z^3} \right] + C
\end{aligned}$$

Substituting to obtain a function of x ,

$$\begin{aligned}
\int \frac{x^2 dx}{(4x+1)^{5/2}} &= \frac{1}{32} \left[\frac{3(4x+1)^2 + 6(4x+1) - 1}{3(4x+1)^{3/2}} \right] + C \\
&= \frac{6x^2 + 6x + 1}{12(4x+1)^{3/2}} + C
\end{aligned}$$

INTEGRATION BY MISCELLANEOUS SUBSTITUTIONS

The substitutions considered in the preceding section made integration possible by rationalizing the function to be integrated. There are also a number of functions which can be integrated by means of substitutions that do not rationalize them. Unfortunately, no general rules can be given for these substitutions and they must be determined by trial and error, guided by experience. A very useful substitution of this type is the *reciprocal substitution*:

$$x = \frac{1}{z} \quad dx = -\frac{dz}{z^2}$$

Algebraic functions which cannot be rationalized frequently can be integrated using this substitution.

EXAMPLES

Evaluate the following integrals.

$$\begin{aligned}
\text{(a) } \int \frac{(x-x^3)^{1/3} dx}{x^4} \\
x &= \frac{1}{z} \\
dx &= -\frac{dz}{z^2} \\
\int \frac{(x-x^3)^{1/3} dx}{x^4} &= -\int \frac{\left(\frac{1}{z} - \frac{1}{z^3}\right)^{1/3} \frac{dz}{z^2}}{\frac{1}{z^4}} \\
&= -\int z(z^2-1)^{1/3} dz \\
&= -\frac{3}{8}(z^2-1)^{4/3} + C
\end{aligned}$$

Substituting to obtain a function of x ,

$$\int \frac{(x - x^3)^{1/3}}{x^4} dx = -\frac{3}{8} \left(\frac{1}{x^2} - 1 \right)^{4/3} + C$$

(b) $\int \frac{dx}{x\sqrt{1+x+x^2}}$ using the standard form

$$\int \frac{du}{\sqrt{a+bu+cu^2}} = \frac{1}{\sqrt{c}} \ln (2cu + b + 2\sqrt{c}\sqrt{a+bu+cu^2}) + C$$

$$x = \frac{1}{z}$$

$$dx = -\frac{dz}{z^2}$$

$$\begin{aligned} \int \frac{dx}{x\sqrt{1+x+x^2}} &= -\int \frac{\frac{dz}{z^2}}{\frac{1}{z}\sqrt{1+\frac{1}{z}+\frac{1}{z^2}}} \\ &= -\int \frac{dz}{z\sqrt{\frac{z^2+z+1}{z^2}}} \\ &= -\int \frac{dz}{\sqrt{z^2+z+1}} \\ &= -\ln (2z+1+2\sqrt{z^2+z+1}) + C \end{aligned}$$

Substituting to obtain a function of x ,

$$\begin{aligned} \int \frac{dx}{x\sqrt{1+x+x^2}} &= -\ln \left(\frac{2}{x} + 1 + 2\sqrt{\frac{1}{x^2} + \frac{1}{x} + 1} \right) + C \\ &= -\ln \left(\frac{2+x}{x} + 2\sqrt{\frac{1+x+x^2}{x^2}} \right) + C \\ &= -\ln \left(\frac{2+x+2\sqrt{1+x+x^2}}{x} \right) + C \\ &= \ln \left(\frac{x}{2+x+2\sqrt{1+x+x^2}} \right) + C \end{aligned}$$

PROBLEMS

Evaluate each of the following integrals by rationalization.

1. $\int \frac{y dy}{\sqrt{2+4y}}$

4. $\int \frac{(\sqrt{x+1}+1) dx}{\sqrt{x+1}-1}$

2. $\int \frac{dt}{2\sqrt{t}+\sqrt[3]{t}}$

5. $\int \frac{dx}{1+\sqrt[3]{x+2}}$

3. $\int \frac{(x^{3/2}-x^{1/3}) dx}{6x^{1/4}}$

6. $\int \frac{dx}{1+\sqrt{x}}$

Evaluate each of the following integrals by reciprocal substitution.

7. $\int \frac{dx}{x^2\sqrt{x^2+a^2}}$

8. $\int \frac{\sqrt{x^2-a^2} dx}{x^4}$

$$9. \int \frac{dx}{x\sqrt{2x-x^2}}$$

$$10. \int \frac{\sqrt{4-x^2} dx}{x^4}$$

$$11. \int \frac{-dx}{x\sqrt{1+4x+5x^2}}$$

$$12. \int \frac{dx}{x^2\sqrt{2x-x^2}}$$

Answers to Odd-Numbered Problems

1. $\frac{1}{6}(2+4y)^{1/2}(y-1) + C$
3. $\frac{2}{27}x^{9/4} - \frac{2}{13}x^{13/12} + C$
5. $\frac{3}{2}(x+a)^{2/3} - 3(x+a)^{1/3} + 3 \ln [(x+a)^{1/3} + 1] + C$
7. $-\frac{\sqrt{x^2+a^2}}{a^2x} + C$
9. $-\left(\frac{2-x}{x}\right)^{1/2} + C$
11. $\ln \left(\frac{1+2x+\sqrt{1+4x+5x^2}}{x} \right) + C$

4.7 Numerical (Approximate) Methods of Integration

Approximate integration is useful in many practical problems when an integral cannot be expressed in terms of elementary functions or when an integrand is defined by an empirical table of values or by an empirical graph. Three methods of approximate integration are discussed in the following sections: the *trapezoidal rule* (linear approximation), *Simpson's rule* (quadratic approximation), and *Taylor's expansion* (series approximation). In all three of these methods, the function to be integrated is replaced by an approximation whose integral is known. The trapezoidal rule and Simpson's rule (but not Taylor's expansion) can also be used when the integrand is defined empirically by a table of values or a graph, rather than by a function.

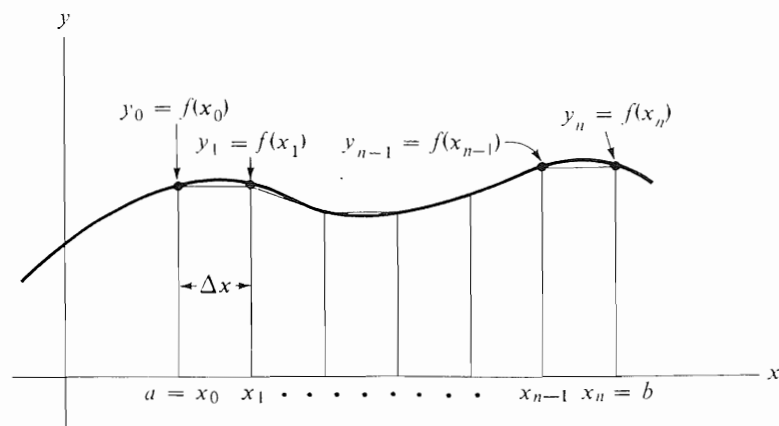
The methods discussed in this section and several other approximation methods are used to evaluate integrals using electronic computers. Approximate integration, even of quite difficult integrals, can be done accurately and efficiently using a computer and, in practice, a computer is used for almost all approximate integration. The main purpose of the following sections is to discuss the principles involved in such approximation. The examples involve fairly crude approximations of relatively simple functions; the amount of calculation required, even for these cases, indicates the advantage of computer approximation.

THE TRAPEZOIDAL RULE

The trapezoidal rule uses a series of straight-line segments to approximate the given function $f(x)$; it is in this sense a method of linear approximation.

Suppose the problem is to evaluate $\int_a^b f(x) dx$. Divide the interval $[a, b]$ into n equal parts, each of length Δx , and denote the points of division by $a = x_0, x_1, \dots, x_n = b$. Erect ordinates at the points of division and denote them by $y_0 = f(x_0), y_1 = f(x_1), \dots, y_n = f(x_n)$. Join the ends of consecutive ordinates to form trapezoids (see Figure 4.31). Then

Figure 4.31



$$\text{Area of first trapezoid} = \frac{1}{2}(y_0 + y_1)\Delta x$$

$$\text{Area of second trapezoid} = \frac{1}{2}(y_1 + y_2)\Delta x$$

$$\vdots$$

$$\text{Area of } n\text{th trapezoid} = \frac{1}{2}(y_{n-1} + y_n)\Delta x$$

and the total area of the n trapezoids is

$$\Delta x \left[\frac{1}{2}y_0 + \sum_{i=1}^{n-1} y_i + \frac{1}{2}y_n \right]$$

The trapezoidal rule for approximate integration is thus

$$\int_a^b f(x) dx \approx \Delta x \left[\frac{1}{2}(y_0 + y_n) + \sum_{i=1}^{n-1} y_i \right]$$

Notes Concerning the Accuracy of the Trapezoidal Rule

1. As the number of divisions n increases, the sum of the areas of the trapezoids more closely approximates the area under the curve.
2. The trapezoidal rule gives a closer approximation to the area under the curve than does the sum of the areas of the corresponding rectangles—that is,

$$\left| \int_a^b f(x) dx - \Delta x \left[\frac{1}{2}(y_0 + y_n) + \sum_{i=1}^{n-1} y_i \right] \right| < \left| \int_a^b f(x) dx - \Delta x \sum_{i=0}^{n-1} y_i \right|$$

and

$$\left| \int_a^b f(x) dx - \Delta x \left[\frac{1}{2}(y_0 + y_n) + \sum_{i=1}^{n-1} y_i \right] \right| < \left| \int_a^b f(x) dx - \Delta x \sum_{i=1}^n y_i \right|$$

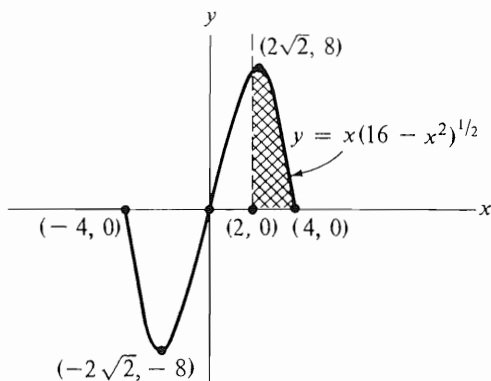
3. The integral $\int_a^b f(x) dx$ is bounded by $\Delta x \sum_{i=0}^{n-1} y_i$ and $\Delta x \sum_{i=1}^n y_i$; depending on which bound is smaller,

$$\Delta x \sum_{i=0}^{n-1} y_i \leq \int_a^b f(x) dx \leq \Delta x \sum_{i=1}^n y_i$$

or

$$\Delta x \sum_{i=1}^n y_i \leq \int_a^b f(x) dx \leq \Delta x \sum_{i=0}^{n-1} y_i$$

The trapezoidal rule averages these two bounds.

EXAMPLEEvaluate $\int_2^4 x(16 - x^2)^{1/2} dx$ by the trapezoidal rule using $n = 4$. (See Figure 4.32).**Figure 4.32**

i	0	1	2	3	4
x_i	2.0	2.5	3.0	3.5	4.0
$f(x_i) = y_i$	$4\sqrt{3} \approx 6.928$	$\frac{5}{4}\sqrt{39} \approx 7.806$	$3\sqrt{7} \approx 7.937$	$\frac{7}{4}\sqrt{15} \approx 6.778$	0

$$\begin{aligned} \int_2^4 x(16 - x^2)^{1/2} dx &\approx \Delta x \left[\frac{1}{2}(y_0 + y_4) + \sum_{i=1}^3 y_i \right] \\ &\approx 0.5 \left[\frac{1}{2}(6.928 + 0) + (7.806 + 7.937 + 6.778) \right] \\ &\approx 0.5 [3.464 + 22.522] \\ &\approx 12.993 \end{aligned}$$

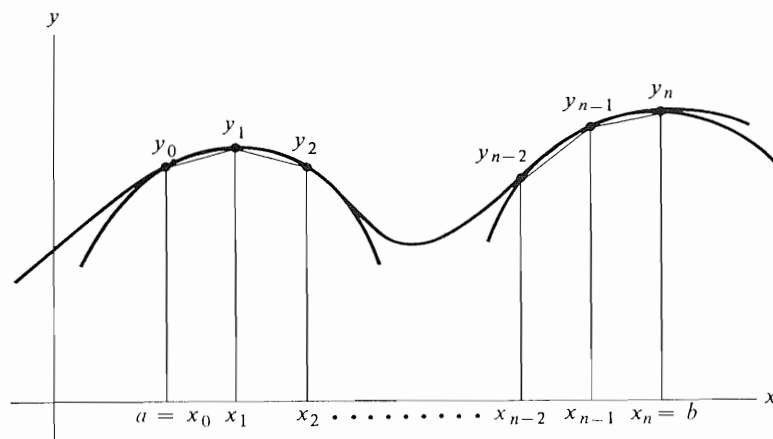
NOTE: $\Delta x \sum_{i=1}^n y_i \leq \int_a^b f(x) dx \leq \Delta x \sum_{i=0}^{n-1} y_i$

$$11.261 \leq \int_2^4 x(16 - x^2)^{1/2} dx \leq 14.725$$

SIMPSON'S RULE

Simpson's rule uses a series of parabolic arcs to approximate the given function $f(x)$; it is in this sense a method of quadratic approximation. Simpson's rule provides a closer approximation than the trapezoidal rule for any given number of subdivisions of the interval over which the integral is to be evaluated.

Figure 4.33



A parabola with vertical axis can be passed through any three points. The equation of this parabola can be obtained from the coordinates of the points; however, for the purpose of using Simpson's rule this is not necessary, since the area under an arc of the parabola can be evaluated without finding the equation of the parabola.

If a parabola with vertical axis is passed through three points (x_0, y_0) , (x_1, y_1) and (x_2, y_2) , where $x_2 - x_1 = x_1 - x_0 = \Delta x$, then the area bounded by the parabola, the x -axis, and the ordinates at x_0 and x_2 is given by

$$A = \frac{\Delta x}{3} (y_0 + 4y_1 + y_2)$$

Again suppose the problem is to evaluate $\int_a^b f(x) dx$. Divide the interval $[a, b]$ into n equal parts, where n must now be an even number, each of length Δx , and denote the points of division by $a = x_0, x_1, \dots, x_n = b$; the corresponding ordinates are y_0, y_1, \dots, y_n ; consider these ordinates in groups of three: y_0, y_1, y_2 ; y_2, y_3, y_4 ; y_4, y_5, y_6 ; \dots ; y_{n-2}, y_{n-1}, y_n , and pass a parabolic arc with vertical axis through each set of three points (see Figure 4.33). The areas under these parabolic arcs are, respectively,

$$A_1 = \frac{\Delta x}{3} (y_0 + 4y_1 + y_2)$$

$$A_2 = \frac{\Delta x}{3} (y_2 + 4y_3 + y_4)$$

$$\vdots$$

$$A_n = \frac{\Delta x}{3} (y_{n-2} + 4y_{n-1} + y_n)$$

and the sum of these areas is

$$\frac{\Delta x}{3} [y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + \dots + 2y_{n-2} + 4y_{n-1} + y_n]$$

Simpson's rule for approximate integration is thus

$$\int_a^b f(x) dx \approx \frac{\Delta x}{3} [y_0 + 4y_1 + 2y_2 + \dots + 2y_{n-2} + 4y_{n-1} + y_n]$$

As in the case of the trapezoidal rule, as the number of divisions, n , increases, the sum of the areas under the parabolic arcs more closely approximates the area under the curve.

EXAMPLE

Evaluate $\int_2^4 x(16 - x^2)^{1/2} dx$ by Simpson's rule using $n = 4$. (See Figure 4.33)

i	0	1	2	3	4
x_i	2.0	2.5	3.0	3.5	4.0
$f(x_i) = y_i$	$4\sqrt{3} \approx 6.928$	$\frac{5}{4}\sqrt{39} \approx 7.806$	$3\sqrt{7} \approx 7.937$	$\frac{7}{4}\sqrt{15} \approx 6.778$	0

$$\begin{aligned}
 \int_2^4 x(16 - x^2)^{1/2} dx &\approx \frac{\Delta x}{3} [y_0 + 4(y_1 + y_3) + 2y_2 + y_4] \\
 &\approx \frac{0.5}{3} [6.928 + 4(7.806 + 6.778) + 2(7.937) + 0] \\
 &\approx \frac{0.5}{3} [6.928 + 58.336 + 15.874] \\
 &\approx 13.523
 \end{aligned}$$

NOTE:

$$\int_2^4 x(16 - x^2)^{1/2} dx = \left[-\frac{1}{3}(16 - x^2)^{3/2} \right]_2^4 = 0 + \frac{1}{3}(12)^{3/2} = 8\sqrt{3} \approx 13.856$$

and, as noted above, Simpson's rule gives a closer approximation than does the trapezoidal rule. Note also that the exact answer is in fact within the bounds given by the trapezoidal rule.

TAYLOR'S EXPANSION

Taylor's theorem may be used to expand a function $f(x)$ in an infinite series; $\int_c^d f(x) dx$ may then be evaluated approximately by integrating the series term by term and evaluating as many terms as necessary to obtain the required accuracy. For this procedure to be valid, the series must converge to $f(x)$ for all values of x in the interval $[c, d]$ over which the integral is to be evaluated. The required accuracy usually can be obtained with fewer terms by expanding the series about a value of a such that $c \leq a \leq d$.

Thus, providing the convergence conditions are met,

$$\begin{aligned}
 \int_c^d f(x) dx &= \int_c^d \left[f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots \right. \\
 &\quad \left. + \frac{f^{(n)}(a)}{n!}(x-a)^n + \cdots \right] dx \\
 &= \left[f(a)x + \frac{f'(a)}{2!}(x-a)^2 + \frac{f''(a)}{3!}(x-a)^3 + \cdots \right. \\
 &\quad \left. + \frac{f^{(n)}(a)}{(n+1)!}(x-a)^{n+1} + \cdots \right]_c^d
 \end{aligned}$$

where, in general, $c \leq a \leq d$.

EXAMPLE

Evaluate $\int_0^{1/2} \frac{x^2 dx}{x+1}$ using Taylor's series.

$$f(x) = \frac{x^2}{x+1}$$

$$f\left(\frac{1}{4}\right) = \frac{\frac{1}{16}}{\frac{5}{4}} = \frac{1}{20}$$

$$f'(x) = \frac{2x(x+1) - x^2}{(x+1)^2} = \frac{x^2 + 2x}{(x+1)^2}$$

$$f'\left(\frac{1}{4}\right) = \frac{\frac{1}{16} + \frac{1}{2}}{\left(\frac{5}{4}\right)^2} = \frac{9}{25}$$

$$f''(x) = \frac{(2x+2)(x+1)^2 - (x^2+2x)(2x+2)}{(x+1)^4}$$

$$f''\left(\frac{1}{4}\right) = \frac{2}{\left(\frac{5}{4}\right)^3} = \frac{128}{125}$$

$$= \frac{2}{(x+1)^3}$$

$$f'''(x) = \frac{-6(x+1)^2}{(x+1)^6} = -\frac{6}{(x+1)^4}$$

$$f'''\left(\frac{1}{4}\right) = -\frac{6}{\left(\frac{5}{4}\right)^4} = -\frac{1536}{625}$$

$$f^{(IV)}(x) = \frac{24(x+1)^3}{(x+1)^8} = \frac{24}{(x+1)^5}$$

$$f^{(IV)}\left(\frac{1}{4}\right) = \frac{24}{\left(\frac{5}{4}\right)^5} = \frac{24576}{3125}$$

$$\vdots$$

$$\vdots$$

$$f^{(n)}(x) = (-1)^n \frac{n!}{(x+1)^{n+1}}$$

$$f^{(n)}\left(\frac{1}{4}\right) = (-1)^n \frac{n!}{\left(\frac{5}{4}\right)^{n+1}}$$

$$\int_c^d f(x) dx = \left[f(a)x + \frac{f'(a)}{2!} (x-a)^2 + \frac{f''(a)}{3!} (x-a)^3 + \cdots \right.$$

$$\left. + \frac{f^{(n)}(a)}{(n+1)!} (x-a)^{n+1} + \cdots \right]_c^d$$

Letting $a = \frac{1}{4}$,

$$\begin{aligned} \int_0^{1/2} \frac{x^2 dx}{x+1} &\approx \left[\frac{1}{20}x + \frac{9}{25 \cdot 2!} \left(x - \frac{1}{4}\right)^2 + \frac{128}{125 \cdot 3!} \left(x - \frac{1}{4}\right)^3 \right. \\ &\quad \left. - \frac{1536}{625 \cdot 4!} \left(x - \frac{1}{4}\right)^4 + \frac{24576}{3125 \cdot 5!} \left(x - \frac{1}{4}\right)^5 \right]_0^{1/2} \\ &\approx \left(\frac{1}{40} + \frac{9}{25 \cdot 2!} \left(\frac{1}{4}\right)^2 + \frac{128}{125 \cdot 3!} \left(\frac{1}{4}\right)^3 - \frac{1536}{625 \cdot 4!} \left(\frac{1}{4}\right)^4 + \frac{24576}{3125 \cdot 5!} \left(\frac{1}{4}\right)^5 \right) \\ &\quad - \left(0 + \frac{9}{25 \cdot 2!} \left(-\frac{1}{4}\right)^2 + \frac{128}{125 \cdot 3!} \left(-\frac{1}{4}\right)^3 \right. \\ &\quad \left. - \frac{1536}{625 \cdot 4!} \left(-\frac{1}{4}\right)^4 + \frac{24576}{3125 \cdot 5!} \left(-\frac{1}{4}\right)^5 \right) \\ &\approx \frac{1}{40} + \frac{2}{375} + \frac{2}{15625} \\ &\approx 0.025 + 0.0053333 + 0.000128 \\ &\approx 0.0304613 \end{aligned}$$

or 0.0305 to four decimal places.

NOTE: Using standard forms,

$$\begin{aligned}
 \int_0^{1/2} \frac{x^2}{x+1} dx &= \int_0^{1/2} \left(x - 1 + \frac{1}{x+1} \right) dx = \left[\frac{x^2}{2} - x + \ln(x+1) \right]_0^{1/2} \\
 &= \left(\frac{1}{8} - \frac{1}{2} + \ln \frac{3}{2} \right) - (0 - 0 + 0) \\
 &= -\frac{3}{8} + \ln \frac{3}{2} \\
 &\approx -0.375 + 0.4055 \\
 &\approx 0.0305 \quad \text{to four decimal places}
 \end{aligned}$$

EXAMPLE

Evaluate $\int_0^1 x^3 dx$ using (a) trapezoidal rule, (b) Simpson's rule, (c) Taylor's series, and (d) exactly.

Divide the range into 10 equal parts as summarized in the following table:

i	0	1	2	3	4	5
x_i	0	0.1	0.2	0.3	0.4	0.5
$f(x_i) = y_i$	0	0.001	0.008	0.027	0.064	0.125
i	6	7	8	9	10	
x_i	0.6	0.7	0.8	0.9	1.0	
$f(x_i) = y_i$	0.216	0.343	0.512	0.729	1.00	

(a) trapezoidal rule: $\int_a^b f(x) dx \approx \Delta x \left[\frac{1}{2}(y_0 + y_n) + \sum_{i=1}^{n-1} y_i \right]$

$$\begin{aligned}
 \int_0^1 x^3 dx &\approx 0.1 \left[\frac{1}{2}(0 + 1.000) + 2.025 \right] \\
 &\approx 0.2525
 \end{aligned}$$

(b) Simpson's rule:

$$\begin{aligned}
 \int_a^b f(x) dx &\approx \frac{\Delta x}{3} [y_0 + 4y_1 + 2y_2 + 4y_3 + \cdots + 4y_{2n-1} + y_n] \\
 \int_0^1 x^3 dx &\approx \frac{0.1}{3} [(0 + 1.000) + 2(0.800) + 4(1.224)] \\
 &\approx 0.2499
 \end{aligned}$$

(c) $\int_c^d f(x) dx = \left[f(a)x + \frac{f'(a)}{2!}(x-a)^2 + \frac{f''(a)}{3!}(x-a)^3 + \cdots \right]_c^d$

Letting $a = \frac{1}{2}$,

$$\int_0^1 x^3 dx = \left[f\left(\frac{1}{2}\right)x + \frac{f'\left(\frac{1}{2}\right)}{2!}\left(x - \frac{1}{2}\right)^2 + \frac{f''\left(\frac{1}{2}\right)}{3!}\left(x - \frac{1}{2}\right)^3 + \cdots \right]_0^1$$

$$f(x) = x^3 \quad f\left(\frac{1}{2}\right) = \frac{1}{8}$$

$$f'(x) = 3x^2 \quad f'\left(\frac{1}{2}\right) = \frac{3}{4}$$

$$f''(x) = 6x \quad f''\left(\frac{1}{2}\right) = 3$$

$$f'''(x) = 6 \quad f'''\left(\frac{1}{2}\right) = 6$$

$$f^{(IV)}(x) = 0 \quad f^{(IV)}\left(\frac{1}{2}\right) = 0$$

$$\begin{aligned}
 \int_0^1 x^3 dx &= \left[\frac{1}{8}x + \frac{\frac{3}{4}}{2!} \left(x - \frac{1}{2}\right)^2 + \frac{3}{3!} \left(x - \frac{1}{2}\right)^3 + \frac{6}{4!} \left(x - \frac{1}{2}\right)^4 \right]_0^1 \\
 &= \left(\frac{1}{8} + \frac{\frac{3}{4}}{2!} \left(\frac{1}{2}\right)^2 + \frac{3}{3!} \left(\frac{1}{2}\right)^3 + \frac{6}{4!} \left(\frac{1}{2}\right)^4 \right) \\
 &\quad - \left(0 + \frac{\frac{3}{4}}{2!} \left(-\frac{1}{2}\right)^2 + \frac{3}{3!} \left(-\frac{1}{2}\right)^3 + \frac{6}{4!} \left(-\frac{1}{2}\right)^4 \right) \\
 &= 0.125 + 0.125 \\
 &= 0.250
 \end{aligned}$$

$$(d) \int_0^1 x^3 dx = \left[\frac{x^4}{4} \right]_0^1 = \frac{1}{4} - 0 = 0.25$$

NOTE: Both Simpson's rule and the Taylor's series expansion provide the exact answer for this particular problem. Because of the nature of the approximation (quadratic), Simpson's rule is exact for any quadratic or cubic function. The Taylor's series expansion is exact if all further terms are zero—that is, if n terms are used and the n th derivative of $f(x)$ is zero. Because Simpson's rule is exact, the answer could have been obtained by dividing the range into only two parts:

i	0	1	2
x_i	0	0.5	1
$f(x_i) = y_i$	0	0.125	1

$$\int_0^1 x^3 dx = \frac{0.5}{3} [1 + 4(0.125)] = 0.25$$

However, using $n = 2$ instead of $n = 10$ would reduce the accuracy of the trapezoidal rule:

$$\int_0^1 x^3 dx = 0.5 \left[\frac{1}{2}(1) + 0.125 \right] = 0.3125$$

EXAMPLE

Evaluate $\int_2^6 \frac{dx}{x^2 - 1}$ using (a) trapezoidal rule, (b) Simpson's rule, (c) Taylor's series, and (d) exactly.

Divide the range into four equal parts as summarized in the following table:

i	0	1	2	3	4
x_i	2	3	4	5	6
$f(x_i) = y_i$	$\frac{1}{3}$	$\frac{1}{8}$	$\frac{1}{15}$	$\frac{1}{24}$	$\frac{1}{35}$
	(0.3333)	(0.1250)	(0.0667)	(0.0417)	(0.0286)

$$(a) \text{ trapezoidal rule: } \int_a^b f(x) dx \approx \Delta x \left[\frac{1}{2}(y_0 + y_n) + \sum_{i=1}^{n-1} y_i \right]$$

$$\int_2^6 \frac{dx}{x^2 - 1} \approx 1 \left[\frac{1}{2}(0.3333 + 0.0286) + 0.2334 \right]$$

$$\approx 0.4143$$

$$(b) \text{ Simpson's rule: } \int_a^b f(x) dx \approx \frac{\Delta x}{3} [y_0 + 4y_1 + 2y_2 + \cdots + 4y_{2n-1} + y_n]$$

$$\int_2^6 \frac{dx}{x^2 - 1} \approx \frac{1}{3} [(0.3333 + 0.0286) + 4(0.1250 + 0.0417) + 2(0.0667)]$$

$$\approx 0.3874$$

$$(c) \int_c^d f(x) dx = \left[f(a)x + \frac{f'(a)}{2!}(x-a)^2 + \frac{f''(a)}{3!}(x-a)^3 + \cdots \right]_c^d$$

Letting $a = 4$,

$$\int_2^6 \frac{dx}{x^2 - 1} \approx \left[f(4)x + \frac{f'(4)}{2!}(x-4)^2 + \frac{f''(4)}{3!}(x-4)^3 + \frac{f'''(4)}{4!}(x-4)^4 + \frac{f^{(iv)}(4)}{5!}(x-4)^5 \right]_2^6$$

$$f(x) = (x^2 - 1)^{-1}$$

$$f'(x) = -1(x^2 - 1)^{-2}(2x)$$

$$= -2x(x^2 - 1)^{-2}$$

$$f''(x) = -2(x^2 - 1)^{-2} - 2x(-2)(x^2 - 1)^{-3}(2x)$$

$$= (x^2 - 1)^{-3}[-2x^2 + 2 + 8x^2]$$

$$= 2(3x^2 + 1)(x^2 - 1)^{-3}$$

$$f'''(x) = 2(6x)(x^2 - 1)^{-3}$$

$$+ 2(3x^2 + 1)(-3)(x^2 - 1)^{-4}(2x)$$

$$= (x^2 - 1)^{-4}[12x^3 - 12x - 36x^3 - 12x]$$

$$= -24x(x^2 + 1)(x^2 - 1)^{-4}$$

$$f^{(iv)}(x) = -24(3x^2 + 1)(x^2 - 1)^{-4} - 24(x^3 + x)(-4)(x^2 - 1)^{-5}(2x)$$

$$= -24(x^2 - 1)^{-5}[3x^4 - 2x^2 - 1 - 8x^4 - 8x^2]$$

$$= 24(5x^4 + 10x^2 + 1)(x^2 - 1)^{-5}$$

$$f(4) = \frac{1}{15}$$

$$f'(4) = \frac{-8}{(15)^2}$$

$$f''(4) = \frac{98}{(15)^3}$$

$$f'''(4) = \frac{-(96)(17)}{(15)^4}$$

$$f^{(iv)}(4) = \frac{(24)(1441)}{(15)^5}$$

$$\begin{aligned}
\int_2^6 \frac{dx}{x^2-1} &\approx \left[\frac{1}{15}x - \frac{8}{(15)^2 \cdot 2!} (x-4)^2 + \frac{98}{(15)^3 \cdot 3!} (x-4)^3 \right. \\
&\quad \left. - \frac{(96)(17)}{(15)^4 \cdot 4!} (x-4)^4 + \frac{(24)(1441)}{(15)^5 \cdot 5!} (x-4)^5 \right]_2^6 \\
&\approx \left(\frac{6}{15} - \frac{8}{(15)^2 \cdot 2!} (2)^2 + \frac{98}{(15)^3 \cdot 3!} (2)^3 - \frac{(96)(17)}{(15)^4 \cdot 4!} (2)^4 + \frac{(24)(1441)}{(15)^5 \cdot 5!} (2)^5 \right) \\
&\quad - \left(\frac{2}{15} - \frac{8}{(15)^2 \cdot 2!} (-2)^2 + \frac{98}{(15)^3 \cdot 3!} (-2)^3 \right. \\
&\quad \left. - \frac{(96)(17)}{(15)^4 \cdot 4!} (-2)^4 + \frac{(24)(1441)}{(15)^5 \cdot 5!} (-2)^5 \right) \\
&\approx 0.2667 + 0.0774 + 0.0243 \\
&\approx 0.3684
\end{aligned}$$

$$\begin{aligned}
\text{(d)} \int_2^6 \frac{dx}{x^2-1} &= \left[\frac{1}{2} \ln \left(\frac{x-1}{x+1} \right) \right]_2^6 = \frac{1}{2} [\ln \frac{5}{7} - \ln \frac{1}{3}] \\
&= \frac{1}{2} [\ln 5 - \ln 7 + \ln 3] \\
&\approx \frac{1}{2} [1.6094 - 1.9459 + 1.0986] \\
&\approx 0.3811
\end{aligned}$$

NOTE: $\int \frac{du}{u^2-a^2} = \frac{1}{2a} \ln \left(\frac{u-a}{u+a} \right) + C.$

It should be noted that the seemingly better results from part (b) are due to the truncation effects of using only 4 terms in (c). Taylor's series is more accurate than Simpson's rule, if a sufficient number of terms are computed.

PROBLEMS

Evaluate each of the following integrals approximately using a. trapezoidal rule, b. Simpson's rule, c. Taylor's series, and d. exactly.

1. $\int_0^3 (x^4 + 3x^2) dx, n = 6$
2. $\int_3^5 \frac{dx}{x}, n = 4$
3. $\int_{-2}^2 e^{2x} dx, n = 4$
4. $\int_{-2}^2 e^{(1/2)x} dx, n = 4$

Evaluate each of the following integrals by expanding the function in a Taylor's series; comment on the interval of convergence and the accuracy of the evaluation.

5. $\int_0^1 (x^4 - x^3 + \frac{1}{2}) dx$
6. $\int_0^2 (5x^4 + \frac{1}{3}x^3 + 6x) dx$
7. $\int_0^2 \left(\frac{x^4}{4} + x^3 + 2x^2 + 1 \right) dx$
8. $\int_2^4 (x^3 - 6x + 3) dx$

Answers to Odd-Numbered Problems

1. a. 78.218, b. 75.625, c. 75.6, d. 75.6
3. a. 35.832, b. 28.904, c. 27.173, d. 27.290
5. $\frac{9}{20}$ (exact, no problem of interval of convergence)
7. $12\frac{1}{3}$ (exact, no problem of interval of convergence)

4.8 Multiple Integrals

Multiple integrals are evaluated by successive partial integration. In this section partial integration, the inverse of partial differentiation, is defined and its use in evaluating multiple integrals is illustrated. The use of double integrals in evaluating probabilities associated with joint-frequency functions is also discussed and illustrated.

The definite integral $\int_a^b f(x) dx$ is defined with respect to the function $f(x)$ over an interval $a \leq x \leq b$. Similarly, the *double integral*

$$\iint_R f(x, y) dy dx$$

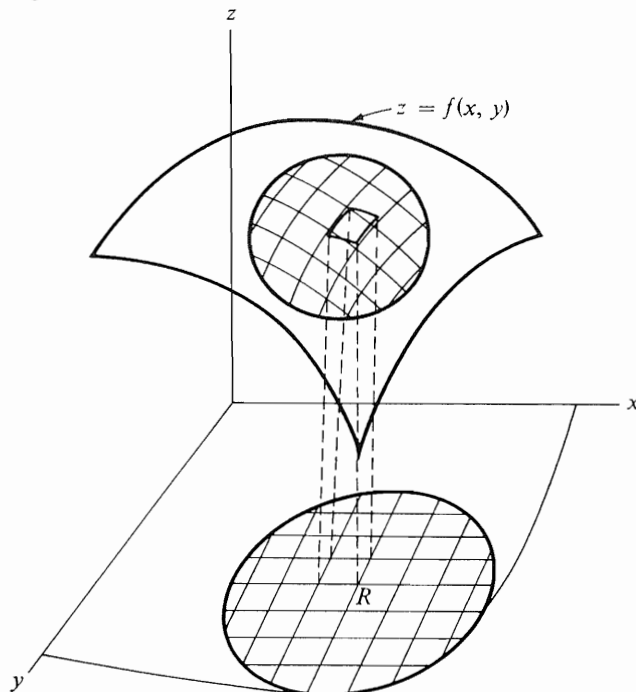
is defined with respect to the function $f(x, y)$ over a bounded region, R , of the xy -plane.

Just as the definite integral of $f(x)$ can be interpreted in terms of area, the definite double integral of $f(x, y)$ can be interpreted in terms of volume. (See Figure 4.34.) The function $f(x, y)$ is assumed to be positive over the region R and the volume computed is that below the surface $z = f(x, y)$ and above the region R in the xy -plane.

The evaluation of double integrals is most easily accomplished by successive *partial integration*, which is the inverse of partial differentiation. That is, to evaluate a double integral, a function of two independent variables is integrated with respect to one of the variables while the other variable is considered as constant; the result of this partial integration is then integrated with respect to the other variable. For this purpose the double integral

$$\iint_R f(x, y) dy dx \quad \text{or} \quad \int_a^b \int_{h(x)}^{g(x)} f(x, y) dy dx$$

Figure 4.34



where a and b are constants, can be written as the *iterated integral*

$$\int_a^b \left[\int_{h(x)}^{g(x)} f(x, y) dy \right] dx$$

To evaluate this expression, $f(x, y)$ is first integrated partially with respect to y and evaluated for the appropriate limits; the result is a function of x which is then integrated with respect to x and evaluated for the appropriate limits. Similarly,

$$\int_a^b \int_{h(y)}^{g(y)} f(x, y) dx dy \quad \text{or} \quad \int_a^b \left[\int_{h(y)}^{g(y)} f(x, y) dx \right] dy$$

is first integrated partially with respect to x and then with respect to y .

NOTE: If the order of integration is reversed, new limits of integration must be determined, that is,

$$\int_a^b \int_{h(x)}^{g(x)} f(x, y) dy dx \neq \int_{h(x)}^{g(x)} \int_a^b f(x, y) dx dy$$

The limits of integration must involve only variables with respect to which integration is to be performed subsequently, and the limits of integration must be rewritten in this manner. Thus

$$\int_a^b \int_{h(x)}^{g(x)} f(x, y) dy dx = \int_c^d \int_{n(y)}^{m(y)} f(x, y) dx dy$$

where the new limits of integration are determined so that, for example, if $y = g(x)$, then $x = m(y)$.

For functions of more than two independent variables, this process can be generalized using multiple integrals; there are relatively few applications requiring more than triple integration. It is important to note that multiple integration is performed from the inside outward; thus the first integral sign belongs to the last differential and so forth. Note also that when an integration is performed with respect to a variable, that variable is eliminated completely from the remaining integral, including the limits of integration.

EXAMPLE

Evaluate the double integral $\int_0^1 \int_{x^2}^x \sqrt{x/y} dy dx$.

$$\begin{aligned} \int_0^1 \int_{x^2}^x \sqrt{\frac{x}{y}} dy dx &= \int_0^1 [2x^{1/2} y^{1/2}]_{x^2}^x dx \\ &= \int_0^1 (2x - 2x^{3/2}) dx \\ &= \left[x^2 - \frac{4x^{5/2}}{5} \right]_0^1 \\ &= (1 - \frac{4}{5}) - 0 = \frac{1}{5} \end{aligned}$$

EXAMPLE

Evaluate the double integral $\int_0^a \int_0^{\sqrt{a^2-y^2}} x \, dx \, dy$.

$$\begin{aligned} \int_0^a \int_0^{\sqrt{a^2-y^2}} x \, dx \, dy &= \int_0^a \left[\frac{x^2}{2} \right]_0^{\sqrt{a^2-y^2}} dy \\ &= \int_0^a \frac{a^2-y^2}{2} dy \\ &= \left[\frac{a^2 y}{2} - \frac{y^3}{6} \right]_0^a \\ &= \left(\frac{a^3}{2} - \frac{a^3}{6} \right) - 0 = \frac{a^3}{3} \end{aligned}$$

EXAMPLE

Evaluate the double integral $\int_0^\pi \int_0^x x \sin y \, dy \, dx$.

$$\begin{aligned} \int_0^\pi \int_0^x x \sin y \, dy \, dx &= \int_0^\pi [-x \cos y]_0^x dx \\ &= \int_0^\pi (-x \cos x + x) dx \\ &= -\int_0^\pi x \cos x \, dx + \left[\frac{x^2}{2} \right]_0^\pi \end{aligned}$$

Integrating $\int_0^\pi x \cos x \, dx$ by parts,

$$\begin{aligned} u &= x & dv &= \cos x \\ du &= dx & v &= \sin x \end{aligned}$$

and

$$\begin{aligned} \int_0^\pi x \cos x &= [x \sin x]_0^\pi - \int_0^\pi \sin x \, dx \\ &= [x \sin x + \cos x]_0^\pi \\ &= (0 - 1) - (0 + 1) \\ &= -2 \end{aligned}$$

Thus

$$\begin{aligned} \int_0^\pi \int_0^x x \sin y \, dy \, dx &= -(-2) + \frac{\pi^2}{2} \\ &= \frac{4 + \pi^2}{2} \end{aligned}$$

EXAMPLE

Evaluate the double integral $\int_{-1}^1 \int_1^{e^y} 1/xy \, dx \, dy$.

$$\begin{aligned} \int_{-1}^1 \int_1^{e^y} \frac{1}{xy} \, dx \, dy &= \int_{-1}^1 \left[\int_1^{e^y} \frac{1}{xy} \, dx \right] dy \\ &= \int_{-1}^1 \left[\frac{1}{y} \ln x \right]_1^{e^y} dy \\ &= \int_{-1}^1 dy \\ &= [y]_{-1}^1 \\ &= 1 - (-1) = 2 \end{aligned}$$

EXAMPLE

Evaluate the triple integral $\int_1^2 y \int_y^{y^2} \int_0^{\ln x} e^z \, dz \, dx \, dy$.

$$\begin{aligned} \int_1^2 y \int_y^{y^2} \int_0^{\ln x} e^z \, dz \, dx \, dy &= \int_1^2 y \int_y^{y^2} [e^z]_0^{\ln x} \, dx \, dy \\ &= \int_1^2 y \int_y^{y^2} (x - 1) \, dx \, dy \\ &= \int_1^2 y \left[\frac{x^2}{2} - x \right]_y^{y^2} dy \\ &= \int_1^2 \left(\frac{y^5}{2} - y^3 - \frac{y^3}{2} + y^2 \right) dy \\ &= \int_1^2 \left(\frac{y^5}{2} - \frac{3y^3}{2} + y^2 \right) dy \\ &= \left[\frac{y^6}{12} - \frac{3y^4}{8} + \frac{y^3}{3} \right]_1^2 \\ &= \left(\frac{16}{3} - 6 + \frac{8}{3} \right) - \left(\frac{1}{12} - \frac{3}{8} + \frac{1}{3} \right) \\ &= \frac{47}{24} \end{aligned}$$

Note that the integrand in the example above can also be written

$$\int_1^2 \int_y^{y^2} \int_0^{\ln x} ye^z \, dz \, dx \, dy$$

and y can be carried as a constant in the first two integrations. However, when an integrand includes a factor which is constant with respect to one or more integrations, that factor can be moved outside the integral(s); usually it is more convenient to separate the integrand in this manner.

EXAMPLE

Evaluate the triple integral $\int_0^\pi \int_0^{\pi/6} \int_0^{a \sin \theta} r^2 \cos \theta \sin \phi \, dr \, d\theta \, d\phi$.

$$\begin{aligned}
\int_0^\pi \int_0^{\pi/6} \int_0^{a \sin \theta} r^2 \cos \theta \sin \phi \, dr \, d\theta \, d\phi &= \int_0^\pi \sin \phi \int_0^{\pi/6} \cos \theta \int_0^{a \sin \theta} r^2 \, dr \, d\theta \, d\phi \\
&= \int_0^\pi \sin \phi \int_0^{\pi/6} \cos \theta \left[\frac{r^3}{3} \right]_0^{a \sin \theta} d\theta \, d\phi \\
&= \int_0^\pi \sin \phi \left[\int_0^{\pi/6} \frac{a^3}{3} \cos \theta \sin^3 \theta \, d\theta \right] d\phi \\
&= \frac{a^3}{3} \int_0^\pi \sin \phi \left[\frac{1}{4} \sin^4 \theta \right]_0^{\pi/6} d\phi \\
&= \frac{a^3}{12} \int_0^\pi \sin \phi \left(\frac{1}{16} \right) d\phi \\
&= \frac{a^3}{192} \int_0^\pi \sin \phi \, d\phi \\
&= \frac{a^3}{192} [-\cos \phi]_0^\pi \\
&= \frac{a^3}{192} (1 + 1) = \frac{a^3}{96}
\end{aligned}$$

Most applications of multiple integration involve definite integrals. This is primarily because any consideration of integration as the reverse process of differentiation is complicated in the case of multiple integration by the form of the constant of integration. For example, when integration is with respect to y , x is constant and, in general, the constant of integration is a function of x . Thus, since multiple integration is the reverse of successive partial differentiations with respect to different variables, the constants of integration are functions of one or more variables; determining these functions requires quite complicated initial conditions. Most problems of this complexity are more conveniently stated in terms of differential equations, as discussed in the next chapter.

PROBLEMS

Evaluate the following multiple integrals.

- $\int_0^1 \int_0^2 (x+2) \, dy \, dx$
- $\int_0^2 \int_1^3 (x+y) \, dx \, dy$
- $\int_0^1 \int_{y^2}^y x^{1/2} \, dx \, dy$
- $\int_0^2 \int_1^{x^2} y \, dy \, dx$
- $\int_0^{-1} \int_{y+1}^{2y} xy \, dx \, dy$
- $\int_0^1 \int_0^{x^2} e^{y/x} \, dy \, dx$
- $\int_b^a \int_0^b \int_a^{2a} x^2 y^2 z \, dz \, dy \, dx$
- $\int_0^1 \int_{y^2}^1 \int_0^{1-x} x \, dz \, dx \, dy$
- $\int_0^1 \int_0^{1-x} \int_0^{1-y^2} z \, dz \, dy \, dx$
- $\int_0^1 \int_0^x \int_0^{x+y} e^{x+y+z} \, dz \, dy \, dx$
- $\int_{-1}^1 \int_y^{2y} x^2 y \, dx \, dy$
- $\int_{-3}^0 \int_0^y y^2 e^{xy} \, dx \, dy$

13. $\int_1^2 \int_0^{x^2} (xy + 1) dy dx$

16. $\int_0^2 \int_y^{y+2} y dx dy$

14. $\int_0^3 \int_x^{x+1} (x + y) dy dx$

17. $\int_0^1 x^2 \int_0^1 e^{xy} dy dx$

15. $\int_{-1}^1 \int_{x-1}^{x+1} (x - y - 1) dy dx$

18. $\int_0^1 \int_0^x \int_0^{x+y} (x + y + z) dz dy dx$

Answers to Odd-Numbered Problems

1. 5

11. $\frac{14}{15}$

3. $\frac{1}{10}$

13. $\frac{91}{12}$

5. $\frac{11}{24}$

15. -4

7. $\frac{1}{6}a^2b^3(a^3 - b^3)$

17. $\frac{1}{2}$

9. $\frac{11}{60}$

PROBABILITY AS A DOUBLE INTEGRAL

One of the most frequent applications of double integrals is in determining the probability that two variables fall within specified ranges. For such applications, the joint-frequency function $f(x, y)$ must be defined.

A joint-frequency function for continuous variables x and y is a function $f(x, y)$ having the following properties:

1. $f(x, y) \geq 0$

2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$

3. $\int_{a_2}^{b_2} \int_{a_1}^{b_1} f(x, y) dx dy = P(a_1 < x < b_1, a_2 < y < b_2)$

These properties are similar to those for a frequency function of one variable and state that (1) probability is always nonnegative; (2) the probability of an event that is certain to occur is one; and (3) the probability that the value of x is in the interval (a_1, b_1) and the value of y is in the interval (a_2, b_2) is given by the corresponding integral. Geometrically, $f(x, y)$ is a surface in three dimensions and the volume under the surface lying above the rectangle determined by $a_1 < x < b_1$ and $a_2 < y < b_2$ is the probability that x and y will assume values corresponding to points in this rectangle. (See Figure 4.29.) If $f(x, y)$ is continuous, $P(a_1 < x < b_1, a_2 < y < b_2) = P(a_1 \leq x \leq b_1, a_2 \leq y \leq b_2)$.

EXAMPLE

In a study of costs for a certain type of fluorescent tube, a company has determined that the frequency function for size of order, x , and total cost for the order, y , is approximately

$$f(x, y) = \frac{1}{3.5} \quad 1 \leq x \leq 6 \quad 0.1 + 0.9x \leq y < 0.1 + 1.1x$$

where x is in thousands of fluorescent tubes and y is in thousands of dollars. If the company sets a price of \$1.05 per tube, on what proportion of the orders will the company break even or make a profit?

$$\begin{aligned}
P(1 \leq x \leq 6, 0.1 + 0.9x \leq y \leq 1.05x) &= \int_1^6 \int_{0.1+0.9x}^{1.05x} \frac{1}{3.5} dy dx \\
&= \frac{1}{3.5} \int_1^6 [y]_{0.1+0.9x}^{1.05x} dx \\
&= \frac{1}{3.5} \int_1^6 (1.05x - 0.1 - 0.9x) dx \\
&= \frac{1}{3.5} \int_1^6 (0.15x - 0.1) dx \\
&= \frac{1}{3.5} [0.075x^2 - 0.1x]_1^6 \\
&= \frac{1}{3.5} [(2.7 - 0.6) - (0.075 - 0.1)] \\
&= \frac{2.125}{3.5} \\
&= 0.6071
\end{aligned}$$

$$\begin{aligned}
\text{NOTE: } \int_1^6 \int_{0.1+0.9x}^{0.1+1.1x} \frac{1}{3.5} dy dx &= \frac{1}{3.5} \int_1^6 [y]_{0.1+0.9x}^{0.1+1.1x} dx \\
&= \frac{1}{3.5} \int_1^6 [(0.1 + 1.1x) - (0.1 + 0.9x)] dx \\
&= \frac{1}{3.5} \int_1^6 0.2x dx \\
&= \frac{1}{3.5} [0.1x^2]_1^6 \\
&= \frac{1}{3.5} [3.6 - 0.1] \\
&= 1
\end{aligned}$$

EXAMPLE

The joint frequency function for the time (in microseconds) required for two switches to close is given by

$$f(x_1, x_2) = 15e^{-(3x_1 + 5x_2)} \quad x_1 > 0, x_2 > 0$$

- Find the probability that the first switch requires less than 0.1 microsecond and the second switch requires more than 0.5 microsecond to close.
- Find the probability that the first switch requires more than 0.2 microsecond and the second switch requires more than 0.5 microsecond to close.

$$\begin{aligned}
 \text{(a) } P(0 < x_1 < 0.1, x_2 > 0.5) &= 15 \int_0^{0.1} \int_{0.5}^x e^{-(3x_1 + 5x_2)} dx_2 dx_1 \\
 &= -3 \int_0^{0.1} [e^{-(3x_1 + 5x_2)}]_{0.5}^{\infty} dx_1 \\
 &= 3 \int_0^{0.1} e^{-(3x_1 + 2.5)} dx_1 \\
 &= [-e^{-(3x_1 + 2.5)}]_0^{0.1} \\
 &= -e^{-2.8} + e^{-2.5} \\
 &= -0.061 + 0.082 \\
 &= 0.021
 \end{aligned}$$

$$\begin{aligned}
 \text{(b) } P(x_1 > 0.2, x_2 > 0.5) &= 15 \int_{0.2}^{\infty} \int_{0.5}^{\infty} e^{-(3x_1 + 5x_2)} dx_2 dx_1 \\
 &= -3 \int_{0.2}^{\infty} [e^{-(3x_1 + 5x_2)}]_{0.5}^{\infty} dx_1 \\
 &= 3 \int_{0.2}^{\infty} e^{-(3x_1 + 2.5)} dx_1 \\
 &= [-e^{-(3x_1 + 2.5)}]_{0.2}^{\infty} \\
 &= 0 + e^{-3.1} \\
 &= 0.045
 \end{aligned}$$

$$\begin{aligned}
 \text{NOTE: } \int_0^{\infty} \int_0^{\infty} 15e^{-(3x_1 + 5x_2)} dx_1 dx_2 &= \int_0^{\infty} [-5e^{-(3x_1 + 5x_2)}]_0^{\infty} dx_2 \\
 &= \int_0^{\infty} 5e^{-5x_2} dx_2 \\
 &= [-e^{-5x_2}]_0^{\infty} \\
 &= 0 - (-1) \\
 &= 1
 \end{aligned}$$

Note also that $f(x_1, x_2) = f(x_1)f(x_2) = (3e^{-3x_1})(5e^{-5x_2})$. Thus

$$P(x_2 > 0.5) = 5 \int_{0.5}^{\infty} e^{-5x_2} dx_2 = [-e^{-5x_2}]_{0.5}^{\infty} = 0 + e^{-2.5} = 0.082$$

$$P(x_1 < 0.1) = 3 \int_0^{0.1} e^{-3x_1} dx_1 = [-e^{-3x_1}]_0^{0.1} = -e^{-0.3} + 1 = 0.259$$

$$P(x_1 > 0.2) = 3 \int_{0.2}^{\infty} e^{-3x_1} dx_1 = [-e^{-3x_1}]_{0.2}^{\infty} = 0 + e^{-0.6} = 0.549$$

These probabilities are referred to as marginal probabilities, and the joint probabilities obtained above can be obtained as the products of the corresponding marginal probabilities. This is characteristic of probabilities of independent events.

EXAMPLE

Show that $f(x, y) = \lambda^2 e^{-\lambda y}$, $0 \leq x \leq y$, is a joint frequency function.

$$f(x, y) \geq 0$$

$$\begin{aligned}\int_0^\infty \int_0^y \lambda^2 e^{-\lambda y} dx dy &= \lambda^2 \int_0^\infty [xe^{-\lambda y}]_0^y dy \\ &= \lambda^2 \int_0^\infty ye^{-\lambda y} dy\end{aligned}$$

Integrating by parts,

$$u = y \quad dv = e^{-\lambda y}$$

$$du = dy \quad v = -\frac{1}{\lambda} e^{-\lambda y}$$

and

$$\begin{aligned}\lambda^2 \int_0^\infty ye^{-\lambda y} dy &= \lambda^2 \left(\left[-\frac{y}{\lambda} e^{-\lambda y} \right]_0^\infty + \frac{1}{\lambda} \int_0^\infty e^{-\lambda y} dy \right) \\ &= \lambda^2 \left[-\frac{y}{\lambda} e^{-\lambda y} - \frac{1}{\lambda^2} e^{-\lambda y} \right]_0^\infty \\ &= [-\lambda y e^{-\lambda y} - e^{-\lambda y}]_0^\infty \\ &= (0 - 0) - (0 - 1) \\ &= 1\end{aligned}$$

PROBLEMS

1. Show that $f(x, y) = \frac{3}{4}(x + y)^2$, $0 \leq x \leq 1$, $0 \leq y \leq 1$ is a frequency function.
2. Show that $f(x, y) = \frac{1}{4}(y - x)^{1/2}$, $0 \leq x \leq y \leq 1$ is a frequency function.
3. Show that $f(x, y) = 8xy$, $0 \leq x \leq 1$, $0 \leq y \leq x$, is a frequency function and find a. $P(0 \leq x \leq 0.5, 0 \leq y \leq x)$ and b. $P(0.2 \leq x \leq 0.5, 0.1 \leq y \leq x)$.
4. Show that $f(x, y) = 12(y - x)^2$, $0 \leq x < y \leq 1$ is a frequency function and find a. $P(0.5 \leq x < y \leq 1)$ and b. $P(0 \leq x < y < 0.5)$.
5. Determine the value of k for which

$$f(x, y) = k(x + y) \quad 0 < x < 4, 0 < y < 6$$

is a frequency function.

6. Determine the value of c for which

$$f(x, y) = \frac{xy}{c} \quad 2 \leq x \leq 4, 5 < y < 9$$

is a frequency function.

7. The joint frequency function for the errors in measuring two variables is given by

$$f(x, y) = \frac{1}{x} \quad 0 < y < x < 1$$

- a. Find the probability that the measurement error x is between 0 and 0.5 and the measurement error y is between 0 and $0.2x$. b. Find the probability that the measurement error x is between 0.8 and 1 and the measurement error y is between 0.8 and 1. c. Show that $f(x, y)$ is a frequency function.
8. The waste from a chemical plant consists of soluble and nonsoluble substances. The nonsoluble substances are of two types, A and B . The plant's engineers have studied data and have determined that the composition of the nonsoluble waste on one day does not depend on its composition on preceding days. The joint frequency function for the proportion of all soluble substances (A and B), denoted by x , and type B nonsoluble substances, denoted by y , is given by

$$f(x, y) = 15xy^2 \quad 0 \leq x \leq 1, 0 \leq y \leq x$$

- a. For what proportion of the days does the waste contain more than 20% of type *B* nonsoluble substances? b. For what proportion of the days does the waste contain less than 50% of nonsoluble substances? c. For what proportion of the days does the waste contain more than 50% nonsoluble substances and more than 20% type *B* substances? d. Show that $f(x, y)$ is a frequency function.

Answers to Odd-Numbered Problems

3. a. 0.0625
b. 0.0567
5. $k = \frac{1}{120}$
7. a. 0.1
b. 0.2

Differential Equations

5

5.1 Introduction

In many problems the relationships between or among variables are most appropriately stated in terms of rates of change. The rate of change of any given variable may be expressed as a function of the rates of change or the values of other variables. For example, in physics equations involving rates of change in energy are used in the derivation of the laws of conservation of energy; in economics it is assumed that the rate at which price approaches its equilibrium value depends on the magnitude of the discrepancy between the quantities supplied and demanded.

Rates of change can be stated in either of two mathematical forms, depending on whether time (or, more generally, any variable with respect to which changes are considered) is viewed as continuous or discrete. When changes are considered to occur continuously or instantaneously, rates of change are stated as derivatives and equations involving them are differential equations. When changes are considered as occurring discretely or discontinuously at certain points in time or as average changes over a period of time, rates of change are stated as differences in the values of variables at different points in time and equations involving them are difference equations. Differential equations are the limiting case of difference equations as the time period between changes or over which an average change is computed approaches zero.

Differential equations, that is, equations involving derivatives, are discussed in this chapter; difference equations, that is, equations involving differences between values of variables at different points in time or with respect to different values of another variable are discussed in Chapter 6.

In this chapter, the definition and classification of differential equations are discussed and types of solutions of ordinary differential equations are considered and illustrated. A classification of differential equations of first order and first degree is given and the solutions of these classes of equations are discussed. Applications of differential equations in economic models are illustrated.

5.2 Definition and Classification of Differential Equations

A *differential equation* is an equation that involves derivatives of one or more functions. Differential equations are classified according to type, order, and degree.

If a differential equation involves derivatives of a function of one independent variable, it is an *ordinary differential equation*; if it involves partial derivatives of a function of two or more independent variables, it is a *partial differential equation*.

The *order* of a differential equation is the order of the highest-order derivative occurring in the equation. The *degree* of a differential equation is the power of the highest-order derivative occurring after the differential equation has been rationalized to remove fractional power derivative terms.

EXAMPLE

$$\frac{dy}{dx} = 2x$$

is an ordinary differential equation of first order and first degree

$$x \, dy - y \, dx = 0$$

is an ordinary differential equation of first order and first degree

$$\frac{d^2 y}{dx^2} + \left(\frac{dy}{dx}\right)^2 + y = 0$$

is an ordinary differential equation of second order and first degree

$$\left(\frac{dy}{dx}\right)^2 = 4 - y^2$$

is an ordinary differential equation of first order and second degree

$$\frac{\partial z}{\partial x} + \frac{\partial z}{\partial y} = z$$

is a partial differential equation of first order and first degree

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + x + y - uz = 0$$

is a partial differential equation of second order and first degree

$$\frac{d^3 z}{dx^3} + x \left(\frac{d^2 z}{dx^2}\right)^2 - xz \frac{dz}{dx} + 10 = 0$$

is an ordinary differential equation of third order and first degree

$$\left(\frac{\partial^3 u}{\partial x^3}\right)^2 + u^2 \left(\frac{\partial u}{\partial x}\right)^3 - x^2 u - 15 = 0$$

is a partial differential equation of third order and second degree

$$\frac{d^4 y}{dx^4} + x^2 \frac{d^2 y}{dx^2} + xy \frac{dy}{dx} - 100y = 0$$

is an ordinary differential equation of fourth order and first degree

$$\frac{\partial^3 y}{\partial x^3} + xy \frac{\partial y}{\partial u} - 3u^4 = 0$$

is a partial differential equation of third order and first degree

$$\left(\frac{dy}{dx}\right)^3 - 5xy^3 + 2 = 0$$

is an ordinary differential equation of first order and third degree

$$\left(\frac{d^4 y}{dx^4}\right)^4 - 6x \left(\frac{d^2 y}{dx^2}\right) - 3y \frac{dy}{dx} = 0$$

is an ordinary differential equation of fourth order and fourth degree

Many of the equations in Chapter 4 are differential equations, since any equation involving at least one derivative is a differential equation. If a differential equation can be written in the form

$$\frac{d^n y}{dx^n} = f(x)$$

its solution is obtained in a straightforward manner by (successive) integration, as in Chapter 4.

EXAMPLE

Find the general solution of the differential equation $\frac{dy}{dx} = \cos x + 2x$.

$$\text{If } \frac{dy}{dx} = \cos x + 2x,$$

$$y = \sin x + x^2 + C$$

EXAMPLE

Find the general solution of the differential equation $\frac{d^2y}{dx^2} = 20x^3 - \frac{1}{x^2}$.

$$\text{If } \frac{d^2y}{dx^2} = 20x^3 - \frac{1}{x^2},$$

$$\frac{dy}{dx} = 5x^4 + \frac{1}{x} + C_1$$

and

$$y = x^5 + \ln x + C_1 x + C_2$$

The methods discussed later in this chapter are appropriate for solving several types of differential equations whose solutions cannot be obtained by straightforward integration. Five types of differential equations of first order and first degree are discussed in some detail; second-order differential equations and first-degree differential equations of higher order are discussed briefly.

5.3 Solutions of Ordinary Differential Equations

A solution of an ordinary differential equation is a function not containing derivatives or differentials, which satisfies the differential equation. Such a solution may be expressed as an explicit or implicit function and may be a general or particular solution, depending on whether the constants of integration have been specified.

The *general solution* of an n th-order differential equation is a solution containing n independent arbitrary constants of integration. For example, the general solution of an n th-order differential equation could be of the implicit form

$$f(x, y) + c_1 + c_2 + \cdots + c_n = 0$$

or the explicit form

$$f(y) = f(x) + c_1 + c_2 + \cdots + c_n$$

A *particular solution* of a differential equation is a solution that can be obtained from the general solution by giving specific values to the arbitrary constants of the general solution. Thus, in either the implicit or explicit form of the solution above, a particular solution is obtained if c_1, c_2, \dots, c_n have specific values.

Some differential equations also have *singular solutions* that cannot be obtained as particular solutions; these exceptional cases are not considered here.

As in simple integration, the constants of integration of a differential equation are specified by *boundary conditions* or *initial conditions*: a condition of the form $y = y_0$ when $x = x_0$ is a boundary condition; the special case $y = y_0$ when $x = 0$ is an initial condition.

Clearly, the problem of the solution or integration of a differential equation is an extension of the problem of simple integration. As in simple integration, frequently a particular solution of a differential equation is needed for practical applications. To obtain such a particular solution it is usually necessary first to find the general solution of the differential equation and then to determine the arbitrary constants from the boundary conditions of the problem. A constant of integration may be written in various forms, such as C , $2C$, C^2 , \sqrt{C} , e^C , $\ln C$; frequently solutions of differential equations may be expressed in simpler form by proper choice of the form of the constant or constants of integration. When the form of a constant of integration is changed, the constant frequently continues to be written using the same notation, with the understanding that it is a "generic constant." Until the numerical value of a constant is determined, its form has no real meaning and thus the constant is written in the simplest way possible.

In each of the following examples, a differential equation and its solution are given and the solution is shown to satisfy the differential equation. Methods for obtaining solutions of differential equations are then discussed.

EXAMPLE

Show that $y = \frac{x^2}{3} + \frac{C}{x}$ is a solution of $x \frac{dy}{dx} + y = x^2$.

$$\text{If } y = x^3/3 + C/x,$$

$$\frac{dy}{dx} = \frac{2x}{3} - \frac{C}{x^2}$$

and thus

$$\underbrace{x \left(\frac{2x}{3} - \frac{C}{x^2} \right) + \frac{x^2}{3} + \frac{C}{x}}_{x \frac{dy}{dx} + y} = x^2$$

EXAMPLE

Show that $y = c_1 e^{kx} + c_2 e^{-kx}$ is a solution of $\frac{d^2 y}{dx^2} - k^2 y = 0$.

$$\text{If } y = c_1 e^{kx} + c_2 e^{-kx},$$

$$\frac{dy}{dx} = c_1 k e^{kx} - c_2 k e^{-kx}$$

$$\frac{d^2 y}{dx^2} = c_1 k^2 e^{kx} + c_2 k^2 e^{-kx}$$

and thus

$$\underbrace{c_1 k^2 e^{kx} + c_2 k^2 e^{-kx} - k^2 (c_1 e^{kx} + c_2 e^{-kx})}_{\frac{d^2 y}{dx^2} - k^2 y} = 0$$

EXAMPLE

Show that $4y = 1/(3x) + c_1 x^5 + c_2 x$ is a solution of $x^2 \frac{d^2 y}{dx^2} - 5x \frac{dy}{dx} + 5y = \frac{1}{x}$.

If $4y = 1/(3x) + c_1 x^5 + c_2 x$,

$$4 \frac{dy}{dx} = -\frac{1}{3x^2} + 5c_1 x^4 + c_2$$

$$4 \frac{d^2 y}{dx^2} = \frac{2}{3x^3} + 20c_1 x^3$$

and thus

$$\frac{1}{4} \left(\frac{2}{3x} + 20c_1 x^5 \right) - \frac{5}{4} \left(-\frac{1}{3x} + 5c_1 x^5 + c_2 x \right) + \frac{5}{4} \left(\frac{1}{3x} + c_1 x^5 + c_2 x \right) = \frac{1}{x}$$

$$x^2 \frac{d^2 y}{dx^2} - 5x \frac{dy}{dx} + 5y$$

$$\underbrace{\left(\frac{2}{12} + \frac{5}{12} + \frac{5}{12} \right) \frac{1}{x}}_{\frac{1}{x}} + \underbrace{\left(5 - \frac{25}{4} + \frac{25}{4} \right) c_1 x^5}_{0} + \underbrace{\left(-\frac{5}{4} + \frac{5}{4} \right) c_2 x}_{0} = \frac{1}{x}$$

EXAMPLE

Show that $y = c_1 e^x + c_2 e^{-2x}$ is a solution of $\frac{d^2 y}{dx^2} + \frac{dy}{dx} - 2y = 0$ and find a particular solution

that satisfies the condition $y = \frac{dy}{dx} = 1$ when $x = 0$.

If $y = c_1 e^x + c_2 e^{-2x}$,

$$\frac{dy}{dx} = c_1 e^x - 2c_2 e^{-2x}$$

$$\frac{d^2 y}{dx^2} = c_1 e^x + 4c_2 e^{-2x}$$

and thus

$$\underbrace{c_1 e^x + 4c_2 e^{-2x} + c_1 e^x - 2c_2 e^{-2x} - 2c_1 e^x - 2c_2 e^{-2x}}_{\frac{d^2 y}{dx^2} + \frac{dy}{dx} - 2y} = 0$$

If $y = y' = 1$ when $x = 0$,

$$c_1 + c_2 = 1$$

$$c_1 - 2c_2 = 1$$

$$c_2 = 0$$

$$c_1 = 1$$

and thus the particular solution is $y = e^x$.

EXAMPLE

Show that $y = 2Cx^2 + C^2$ is a solution of $\left(\frac{dy}{dx}\right)^2 + 8x^3 \frac{dy}{dx} = 16x^2y$ and find a particular solution that satisfies the condition $y = -1$ when $x = 1$.

$$\text{If } y = 2Cx^2 + C^2,$$

$$\frac{dy}{dx} = 4Cx$$

and thus

$$\underbrace{16C^2x^2 + 8x^3(4Cx)}_{\left(\frac{dy}{dx}\right)^2 + 8x^3\left(\frac{dy}{dx}\right)} = \underbrace{16x^2(2Cx^2 + C^2)}_{16x^2y}$$

If $y = -1$ when $x = 1$,

$$-1 = 2C + C^2$$

$$C^2 + 2C + 1 = 0$$

$$(C + 1)^2 = 0$$

$$C = -1,$$

and thus the particular solution is $y = 1 - 2x^2$.

EXAMPLE

Show that $(x - c_1)^2 + y^2 = c_2$ is a solution of

$$y \frac{d^2y}{dx^2} + \left(\frac{dy}{dx}\right)^2 + 1 = 0$$

and find a particular solution that satisfies the condition $y = 3$, and $\frac{dy}{dx} = -\frac{4}{3}$ when $x = 5$.

$$\text{If } (x - c_1)^2 + y^2 = c_2,$$

$$2y \frac{dy}{dx} = -2(x - c_1)$$

$$\frac{dy}{dx} = -\frac{x - c_1}{\sqrt{c_2 - (x - c_1)^2}}$$

$$\frac{d^2y}{dx^2} = \frac{-\sqrt{c_2 - (x - c_1)^2} + (x - c_1)\left(\frac{1}{2}\right)[c_2 - (x - c_1)^2]^{-1/2}(-2)(x - c_1)}{c_2 - (x - c_1)^2}$$

$$= \frac{-1}{\sqrt{c_2 - (x - c_1)^2}} - \frac{(x - c_1)^2}{[c_2 - (x - c_1)^2]^{3/2}}$$

and thus

$$-1 - \underbrace{\frac{(x - c_1)^2}{c_2 - (x - c_1)^2}}_{y \frac{d^2 y}{dx^2}} + \underbrace{\left[-\frac{x - c_1}{\sqrt{c_2 - (x - c_1)^2}} \right]^2}_{\left(\frac{dy}{dx} \right)^2} + 1 = 0$$

If $y = 3$ and $\frac{dy}{dx} = -\frac{4}{3}$ when $x = 5$,

$$3 = \sqrt{c_2 - (5 - c_1)^2}$$

$$\frac{4}{3} = \frac{5 - c_1}{\sqrt{c_2 - (5 - c_1)^2}}$$

$$5 - c_1 = 4$$

$$c_1 = 1$$

$$c_2 = 25$$

and thus the particular solution is $(x - 1)^2 + y^2 = 25$.

PROBLEMS

State the type, order, and degree of each of the following differential equations.

- $\frac{d^2 y}{dx^2} + 6x + y = 0$
- $x^2 dy = xy^2 dx$
- $\left(\frac{d^2 y}{dx^2} \right)^2 + 5xy = 0$
- $\frac{\partial z}{\partial x} + \frac{\partial z}{\partial y} = xyz$
- $\left(\frac{\partial z}{\partial x} \right)^2 + \frac{\partial z}{\partial y} - 2x = 0$
- $3x \left(\frac{dy}{dx} \right)^2 = 4xy$
- $\frac{\partial^2 z}{\partial x^2} = \frac{\partial z}{\partial y}$
- $\frac{d^2 y}{dx^2} + \frac{dy}{dx} - 3x = 0$
- $\left(\frac{d^2 y}{dx^2} \right)^2 - \frac{dy}{dx} = 4y$
- $\frac{d^3 y}{dx^3} = 10y$
- Verify that $y^2 = Cx + \frac{1}{8}C^3$ is a solution of $y = 2x \frac{dy}{dx} + y^2 \left(\frac{dy}{dx} \right)^3$ and find the particular solution if $y' = 1$ when $x = 0$.
- Verify that $(x - c_1)^2 + y^2 = c_2$ is a solution of $y \frac{d^2 y}{dx^2} + \left(\frac{dy}{dx} \right)^2 + 1 = 0$ and find the particular solution if $y = 3$ and $y' = -\frac{4}{3}$ when $x = 5$.
- Verify that $\ln y = c_1 e^x + c_2 e^{-x} + x^2 + 2$ is a solution of $\left(\frac{1}{y} \frac{dy}{dx} \right)^2 - \frac{1}{y} \frac{d^2 y}{dx^2} = x^2 - \ln y$.
- Verify that $x = \cos 2t + 2c_1 \cos 3t + 3c_2 \sin 3t$ is a solution of $\frac{d^2 x}{dt^2} + 9x = 5 \cos 2t$.
- Verify that $y = (c_1 + c_2 \ln x)\sqrt{x} + c_3$ is a solution of $4x^2 \frac{d^3 y}{dx^3} + 8x \frac{d^2 y}{dx^2} + \frac{dy}{dx} = 0$.

16. Verify that $y = x^3 + c_1 x^2 + c_2$ is a solution of $\frac{d^2 y}{dx^2} - \frac{1}{x} \frac{dy}{dx} - 3x = 0$ and find the particular solution if $y = 1$ when $x = 1$ and $y = 5$ when $x = 2$.
17. Verify that $y = c_1 \cos x + c_2 \sin x - \frac{1}{3} \cos 2x$ is a solution of $\frac{d^2 y}{dx^2} + y = \cos 2x$.
18. Verify that $y = e^{-x}(x + C)$ is a solution of $\frac{dy}{dx} + y = e^{-x}$.
19. Verify that $y = c_1 x + \frac{c_2}{x} + c_3$ is a solution of $\frac{d^3 y}{dx^3} + \frac{3}{x} \frac{d^2 y}{dx^2} = 0$.
20. Verify that $y^2 = c_1 x^2 + c_2 x$ is a solution of $x^2 y \frac{d^2 y}{dx^2} + \left(x \frac{dy}{dx} - y \right)^2 = 0$ and find a particular solution that satisfies the condition $y = 2$ and $\frac{dy}{dx} = 1$ when $x = 1$.

Answers to Odd-Numbered Problems

- | | |
|--|--|
| 1. ordinary, second order, first degree | 7. partial, second order, first degree |
| 3. ordinary, second order, second degree | 9. ordinary, second order, second degree |
| 5. partial, first order, second degree | 11. $y^2 = 2x + 1$ |

5.4 Differential Equations of First Order and First Degree

In this section differential equations of first order and first degree are classified as separable, homogeneous, exact, linear, or linear in a function. Methods of solution for these types of equations are discussed and illustrated.

A differential equation of first order and first degree can be written in the form

$$\frac{dy}{dx} = F(x, y)$$

If $F(x, y)$ is a constant or a function only of x , the differential equation is solved by the usual methods of integration; if $F(x, y)$ is in fact a function of both x and y , the methods discussed below are appropriate for solving the differential equation.

An alternative form, frequently more useful for classifying differential equations of the first order and first degree, is

$$M(x, y) dx + N(x, y) dy = 0$$

There is one difference between these equations which should be noted. In the first equation y is the dependent variable and x is the independent variable; thus the solution expresses y as a function of x and an arbitrary constant. In the second equation the relationship is expressed implicitly and the choice of the dependent and independent variables is a matter of convenience; the solution is frequently also expressed as an implicit function.

Methods for the solution of the following types of differential equations of the first order and first degree are discussed in considerable detail; note that facility in solving a differential equation of any order and degree depends to a very considerable extent on classifying it correctly.

1. *Separable differential equations*: M is a function only of x and N is a function only of y , so the equation is of the form $M(x) dx + N(y) dy = 0$.
2. *Homogeneous differential equations*: M and N are homogeneous functions of the same degree of homogeneity.
3. *Exact differential equations*: $\frac{\partial}{\partial y} (M(x, y)) = \frac{\partial}{\partial x} (N(x, y))$ so the equation is of the form $\frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy = 0$, where $F(x, y)$ is the solution of the differential equation.
4. *Linear differential equations*: The equation is of the first degree in y and $\frac{dy}{dx}$ or in x and $\frac{dx}{dy}$ and thus is of the form $\frac{dy}{dx} + yP(x) = Q(x)$ or $\frac{dx}{dy} + xP(y) = Q(y)$.
5. *Differential equations linear in a function of y or in a function of x* : The equation is of the first degree in $f(y)$ and $\frac{d}{dy} f(y)$ or in x and $\frac{d}{dx} f(x)$ and thus is of the form $\frac{d}{dy} f(y) + f(y)P(x) = Q(x)$ or $\frac{d}{dx} f(x) - f(x)P(y) = Q(y)$.

SEPARABLE DIFFERENTIAL EQUATIONS

If a differential equation can be written in the form

$$M(x) dx + N(y) dy = 0$$

where, as indicated, M is a function only of x and N is a function only of y , then the variables are said to be *separated* and the general solution of the differential equation is obtained by the usual methods of integration:

$$\int M(x) dx + \int N(y) dy = C$$

and

$$f(x) + g(y) = C$$

EXAMPLE

Solve the equation $(1 + x^2) \frac{dy}{dx} + xy = 0$.

$$(1 + x^2) dy + xy dx = 0$$

$$\frac{x}{1 + x^2} dx + \frac{1}{y} dy = 0$$

The variables are separated and, integrating,

$$\frac{1}{2} \ln (1 + x^2) + \ln y = C$$

$$\ln (y\sqrt{1 + x^2}) = C$$

$$y\sqrt{1 + x^2} = C$$

Note that the constant in the last equation is actually e^C in terms of the preceding equations. As mentioned in the discussion of constants of integration above, until a constant is evaluated to obtain a particular solution, its form is arbitrary and the constant is thus written in the most convenient way.

EXAMPLE

Solve the equation $y^2 dx - (1 - x) dy = 0$.

$$\frac{1}{x-1} dx + \frac{1}{y^2} dy = 0$$

The variables are separated and, integrating,

$$\ln(x-1) - \frac{1}{y} = C$$

$$y \ln(x-1) - 1 = Cy$$

$$y \ln[C(x-1)] = 1$$

or

$$C(x-1) = e^{1/y}$$

EXAMPLE

If the interest rate is 100*i*% compounded continuously and A is the amount at any time (principal plus any accumulated interest), then

$$\frac{dA}{dt} = iA$$

and, separating the variables,

$$\frac{dA}{A} = i dt$$

and, integrating,

$$\int \frac{dA}{A} = \int i dt$$

$$\ln A = it + c$$

$$A = e^{it+c}$$

or

$$A = ce^{it}$$

If $A = A_0$ at $t = 0$, then $c = A_0$ and

$$A = A_0 e^{it}$$

NOTE: The formula for discrete compounding is given by

$$y = x \left(1 + \frac{i}{k} \right)^{nk}$$

where y is the amount after n years (principal plus any accumulated interest) and interest rate is $i\%$ per year compounded k times per year. In the notation of this section, the discrete formula can be written

$$A = A_0 \left(1 + \frac{i}{k} \right)^{kt}$$

and the formula for the continuous case can be obtained by taking the limit as $k \rightarrow \infty$ of the formula for the discrete case, since

$$e = \lim_{k \rightarrow \infty} \left(1 + \frac{1}{k} \right)^k$$

EXAMPLE

The relationship between net profit P and advertising expenditure x is such that the rate of increase of net profit as advertising expenditure increases is proportional to a constant minus net profit. Find the relationship between net profit and advertising expenditure if $P = P_0$ when $x = 0$.

$$\frac{dP}{dx} = k(a - P)$$

$$\frac{dP}{a - P} = k dx$$

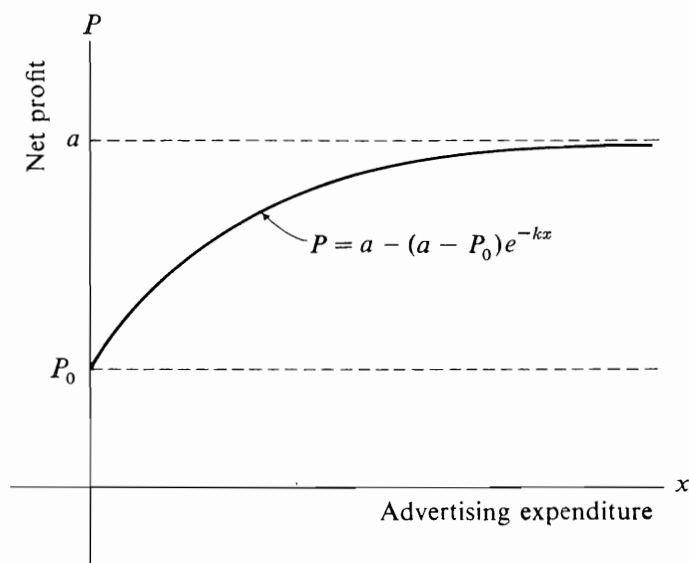
$$-\ln(a - P) = kx + C$$

$$\ln(a - P) = -kx + C$$

$$a - P = Ce^{-kx}$$

$$P = a - Ce^{-kx}$$

Figure 5.1



so

$$P_0 = a - C$$

$$C = a - P_0$$

and

$$P = a - (a - P_0)e^{-kx}$$

Thus net profit is P_0 with no advertising expenditure and increases, as advertising expenditure increases, toward an (asymptotic) maximum of a . (See Figure 5.1.)

HOMOGENEOUS DIFFERENTIAL EQUATIONS

A differential equation of the form

$$M(x, y) dx + N(x, y) dy = 0$$

is said to be homogeneous if $M(x, y)$ and $N(x, y)$ are homogeneous functions of the same degree in x and y . Recall that $F(x, y)$ is a homogeneous function of degree n in x and y if and only if $F(kx, ky) = k^n F(x, y)$, where k is any constant. When a differential equation is homogeneous, its variables can be separated by the substitution

$$y = vx$$

$$dy = v dx + x dv$$

or, equivalently, by the substitution

$$x = vy$$

$$dx = v dy + y dv$$

The resulting differential equations can then be written, respectively, in the form

$$M(x) dx + N(v) dv = 0$$

or in the form

$$M(v) dv + N(y) dy = 0$$

and solved by the usual methods of integration. The general solution of the original equation is then obtained by substituting, respectively,

$$v = \frac{y}{x} \quad \text{or} \quad v = \frac{x}{y}$$

in the solution of the separable differential equation.

NOTE: This procedure can be justified mathematically as follows. If $M(x, y) dx + N(x, y) dy = 0$, $M(x, y)$ and $N(x, y)$ are homogeneous of the same degree n and $N(x, y) \neq 0$, then

$$\frac{dy}{dx} = -\frac{M(x, y)}{N(x, y)}$$

But, since $M(x, y) = t^{-n}M(tx, ty)$ and $N(x, y) = t^{-n}N(tx, ty)$,

$$\frac{dy}{dx} = -\frac{M(tx, ty)}{N(tx, ty)}$$

for any value of t . If $t = 1/x$, then

$$\frac{dy}{dx} = -\frac{M\left(1, \frac{y}{x}\right)}{N\left(1, \frac{y}{x}\right)}$$

and the differential equation can be written

$$\frac{dy}{dx} = F\left(\frac{y}{x}\right)$$

Substituting $y = vx$ and $\frac{dy}{dx} = v + x \frac{dv}{dx}$, this equation can be written

$$v + x \frac{dv}{dx} = F(v)$$

$$(v - F(v)) dx + x dv = 0$$

$$\frac{dx}{x} + \frac{dv}{v - F(v)} = 0$$

and the variables x and v are separated. If $v - F(v) = 0$, then $x dv = 0$, $dy - v dx = 0$, and $M(x, y) dx + N(x, y) dy = 0$ has the simple form $y dx - x dy = 0$.

A similar argument shows that the substitution $x = vy$ also results in an equation in which the variables are separable.

EXAMPLE

Solve the equation $(y^2 - xy) dx + x^2 dy = 0$.

The equation is homogeneous (degree 2). Substituting $y = vx$, $dy = v dx + x dv$,

$$(v^2x^2 - vx^2) dx + x^2(v dx + x dv) = 0$$

$$v^2x^2 dx + x^3 dv = 0$$

$$\frac{1}{x} dx + \frac{1}{v^2} dv = 0$$

The variables are separated and, integrating,

$$\ln x - \frac{1}{v} = C$$

Substituting $v = y/x$,

$$\ln x - \frac{x}{y} = C$$

$$y = \frac{x}{\ln x - C}$$

EXAMPLE

Solve the equation $(x + y) dx + (x - y) dy = 0$.

The equation is homogeneous (degree 1). Substituting $x = vy$, $dx = v dy + y dv$,

$$(vy + y)(v dy + y dv) + (vy - y) dy = 0$$

$$(vy^2 + y^2) dv + (v^2y + 2vy - y) dy = 0$$

$$y^2(v + 1) dv + y(v^2 + 2v - 1) dy = 0$$

$$\frac{v + 1}{v^2 + 2v - 1} dv + \frac{1}{y} dy = 0$$

The variables are separated and, integrating,

$$\frac{1}{2} \ln(v^2 + 2v - 1) + \ln y = C$$

$$y(v^2 + 2v - 1)^{1/2} = C$$

$$y^2(v^2 + 2v - 1) = C$$

Substituting $v = x/y$,

$$x^2 + 2xy - y^2 = C$$

EXAMPLE

The relationship between the manufacturing cost per item, M , and the number of types of items manufactured, N , is such that the rate of increase of manufacturing cost as number of types of items increases is equal to the ratio of the cost per item plus the number of types of items divided by the number of types of items. Find the relationship between manufacturing cost per item and number of types of items manufactured if $M = M_0$ when $N = 1$.

$$\frac{dM}{dN} = \frac{M + N}{N}$$

$$N dM = (M + N) dN$$

$$M = vN$$

$$dM = v dN + N dv$$

$$vN dN + N^2 dv = vN dN + N dN$$

$$dv = \frac{dN}{N}$$

$$v = \ln N + C$$

Substituting $v = M/N$,

$$\frac{M}{N} = \ln N + C$$

$$M = N \ln N + NC$$

so

$$M_0 = 0 + C$$

$$C = M_0$$

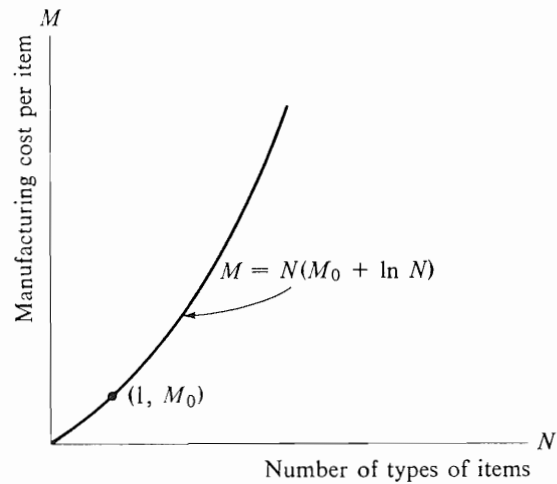
and

$$M = N \ln N + NM_0$$

$$M = N(M_0 + \ln N)$$

(see Figure 5.2).

Figure 5.2



EXAMPLE

Suppose that the rate of increase in the cost of ordering and holding, y , as the size of the order, s , increases is equal to the ratio of the sum of the squares of the cost and the size divided by twice the product of the cost and the size. Find the relationship between the cost of ordering and holding and the size of the order if $y = 3$ when $s = 1$.

$$\frac{dy}{ds} = \frac{y^2 + s^2}{2sy}$$

$$2sy \, dy - (y^2 + s^2) \, ds = 0$$

$$y = vs$$

$$dy = v \, ds + s \, dv$$

$$2vs^2(v \, ds + s \, dv) - (v^2s^2 + s^2) \, ds = 0$$

$$(v^2s^2 - s^2) \, ds + 2vs^3 \, dv = 0$$

$$s^2(v - 1) \, ds + 2vs^3 \, dv = 0$$

$$\frac{ds}{s} + \frac{2v}{v^2 - 1} \, dv = 0$$

$$\ln s + \ln(v^2 - 1) = C$$

$$s(v^2 - 1) = C$$

Substituting $v = y/s$,

$$s\left(\frac{y^2}{s^2} - 1\right) = C$$

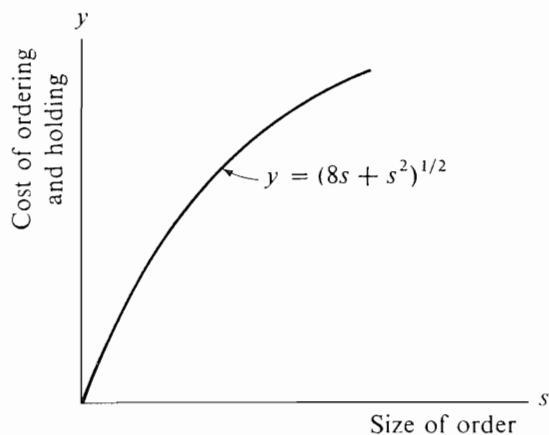
$$y^2 - s^2 = Cs$$

$$y = (Cs + s^2)^{1/2}$$

so

$$3 = (C + 1)^{1/2}$$

Figure 5.3



and

$$C = 8$$

$$y = (8s + s^2)^{1/2}$$

(see Figure 5.3).

PROBLEMS

Find the general solution for each of the following differential equations.

1. $(x^2 + y^2) dx - 2xy dy = 0$
2. $(xy - x^2) \frac{dy}{dx} = y^2$
3. $y^3 dx - x^3 dy = 0$
4. $(y + 3) dx + \cot x dy = 0$
5. $\frac{dx}{dt} = 1 - \sin 2t$
6. $\frac{d^3 y}{dx^3} = e^{-x}$
7. $\frac{d^2 y}{dx^2} = \frac{1}{x^2}$
8. $\sec x \cos^2 y dx = \cos x \sin y dy$
9. $\sin x \cos^2 y dx + \cos^2 x dy = 0$
10. $dr + r \cos \theta d\theta = 0$
11. $\frac{dy}{dx} = \frac{x - y}{x + y}$
12. $(xy^2 - x) dx + (x^2 y + y) dy = 0$
13. $\frac{dy}{dx} = \frac{xy + y}{x + xy}$
14. $\frac{dy}{dx} = \frac{2x - y}{x + 4y}$
15. $dy = 2xy dx$
16. $\frac{1}{\rho} \frac{d\rho}{d\theta} = \frac{\rho^2 - 1}{\rho^2 + 1} \tan \theta$
17. $r d\theta + \theta dr = 2 dr$
18. $d\rho + \rho \tan \theta d\theta = 0$
19. $(r^2 + 1) d\theta + \frac{dr}{\sec^2 \theta} = 0$
20. $(xy - y^2) dx + \left(y^2 - \frac{x^2}{2}\right) dy = 0$
21. $\frac{y^2}{x} dx + (x - y) dy = 0$
22. $(xy^3 - x) dx + xy^2 dy = 0$
23. $\tan x \sin^2 y dx + \cos^2 x \cot y dy = 0$
24. $x \frac{dy}{dx} - y - x \sin \frac{y}{x} = 0$
25. $\sin x \cos y dx + \cos x \sin y dy = 0$

26. $\frac{dy}{dx} = -\frac{y}{x}$

31. $(x^3 - y^3) dx + xy^2 dy = 0$

32. $(x + y) dx + x dy = 0$

27. $x^3 - 2y^3 + 3xy^2 \frac{dy}{dx} = 0$

33. $\frac{dy}{dx} = e^{x-y}$

28. $\frac{dy}{dx} = \frac{x^2}{y^3}$

34. $x(2y - 3) dx + (x^2 + 1) dy = 0$

29. $x \, dy - y \, dx = 0$

35. $(x + y) dx = x dy$

30. $(x^2 + 2y^2) dx = xy dy$

36. $x \, dy - y \, dx = \sqrt{xy} \, dx$

Find the particular solution for each of the following differential equations under the given conditions.

37. $\frac{dy}{dx} = \frac{y}{x}$, $y = 3$ when $x = 1$

38. $xy \, dx + \sqrt{1+x^2} \, dy = 0$, $y = 1$ when $x = 0$

39. $(x^2 + y^2) dx = 2xy dy$, $y = 0$ when $x = 1$

40. $x \, dy + 2y \, dx = 0$, $y = 1$ when $x = 2$

41. $(xe^{y/x} + y) dx = x dy$, $y = 0$ when $x = 1$

42. $y^2(y \, dx - x \, dy) + x^3 \, dx = 0$, $y = 3$ when $x = 1$

43. $x \, dx - 4y \, dy = 0$, $y = 2$ when $x = 5$

44. $x(y + 1) dx + y(x + 1) dy = 0$, $y = 1$ when $x = 0$

45. $(x - \sqrt{xy}) dy = y dx$, $y = 1$ when $x = 4$

46. $(x + 4)(2y + 6) dy + xy^2 dx = 0$, $y = 16$ when $x = 0$

47. $(xy^2 + x^2y) dy - xy^2 dx = 0$, $y = 1$ when $x = 6$

48. $y(x^2 + 6) dy + x(y^2 + 1) dx$, $y = 3$ when $x = 0$

49. $\frac{1}{xy} dx + \frac{e^{y^2}}{x^3 - 1} dy = 0$, $y = 0$ when $x = 1$

50. $\frac{dy}{dx} = \frac{y}{x} + \tan \frac{y}{x}$, $y = \pi$ when $x = 6$

51. $\cot y \, dx + \cot x \, dy = 0$, $y = 0$ when $x = 0$

52. The relationship between price, p , and quantity demanded, x , is such that the rate of decrease in demand as price increases is proportional to the quantity demanded and inversely proportional to the price plus a constant. Find the demand function if $p = p_0$ when $x = 1$. Sketch the relationship obtained.

53. The rate of increase of total cost, y , as number of units manufactured, x , increases is proportional to the number of units manufactured plus a constant and inversely proportional to the total cost. Find the cost function if $y = y_0$ when $x = 0$. Sketch the relationship obtained.

54. The rate of increase in sales, s , as advertising effort, x , increases is equal to a constant plus advertising effort. Find the relationship between sales and advertising effort if $s = s_0$ when $x = 0$. Sketch the relationship obtained.

55. The relationship between revenue, R , and quantity demanded, x , is such that the rate of increase in revenue as quantity demanded increases is equal to twice the cube of the revenue minus the cube of the quantity demanded, all divided by three times the product of the quantity demanded and the square of the revenue. Find the relationship between revenue and quantity demanded if $R = 0$ when $x = 10$. Sketch the relationship obtained.

56. The relationship between average cost, \bar{y} , and number of units produced, x , is such that the change in average cost as the number of units increases is equal to the ratio of the number of units minus average cost divided by number of units. Find the relationship between average cost and the number of units produced if $\bar{y} = \frac{2}{3}$ when $x = 1$. Sketch the relationship obtained.
57. The rate of increase in cost, y , as the number of units manufactured, x , increases is equal to the ratio of twice the square of the cost minus the square of the number of units divided by the product of the cost and the number of units. Find the relationship between cost and the number of units manufactured if $y = 3$ when $x = 1$. Sketch the relationship obtained.
58. The rate of increase of sales volume, s , as price, p , decreases is proportional to sales volume and inversely proportional to price minus a constant. Find the relationship between sales volume and price if $s = s_0$ when $p = p_0$.
59. If interest is compounded continuously, a. find the amount available after 10 years if \$5000 is deposited at 4%; and b. find the amount available after 20 years if \$20,000 is deposited at 6%. (Compare the answers with those obtained for the example on page 118 and Problem 2 on page 123, respectively.)
60. If growth is continuous and the rate, r , is proportional to the number, N , present in the population, then

$$\frac{dN}{dt} = rN$$

If $N = N_0$ when $t = 0$, obtain a formula for the number in the population at time t . (Compare with the discrete case on page 119.)

61. If the growth rate is continuous at the rate of 5% per year and the original number in the population is 200, what is the size of the population after 10 years?
62. A manufacturer has found that the change in the cost of distribution, D , as sales, S , increases is equal to a constant times sales plus another constant. Find the relationship between the cost of distribution and sales if $D = 0$ when $S = 0$. Sketch the relationship obtained.
63. The rent for an apartment (two-bedroom, standard furnishings) in a college town varies with the distance of the apartment from campus. Suppose this relationship is given by

$$\frac{dy}{dx} = -\left(\frac{k}{x} + a\right) \quad 1 \leq x \leq 10$$

where y is monthly rent (in dollars) and x is distance (in miles) and k and a are constants. If $y = 225$ when $x = 1$, find y as a function of x and sketch the relationship obtained.

64. The relationship between the cost of operating a warehouse and the number of gallons of oil stored in the warehouse is given by

$$\frac{dy}{dx} = kx + a$$

where y is the monthly cost of operating the warehouse (in dollars) and x is the number of gallons of oil in storage. If $y = y_0$ (fixed cost) when $x = 0$, find y as a function of x and sketch the relationship obtained.

Answers to Odd-Numbered Problems

- | | |
|--------------------------------------|--|
| 1. $x^2 - y^2 = Cx$ | 13. $y - x + \ln \frac{y}{x} = C$ |
| 3. $x^2 - y^2 = Cx^2y^2$ | 15. $y = Ce^{x^2}$ |
| 5. $x = t + \frac{1}{2} \cos 2t + C$ | 17. $r(\theta - 2) = C$ |
| 7. $y = -\ln x + c_1x + c_2$ | 19. $\tan \theta + \ln \sqrt{r^2 + 1} = C$ |
| 9. $\sec x + \tan y = C$ | 21. $y = Ce^{y/x}$ |
| 11. $y^2 + 2xy - x^2 = C$ | 23. $\tan^2 x - \cot^2 y = C$ |

25. $\cos y = C \sec x$ 47. $\ln y - \frac{x}{y} + 6 = 0$
 27. $x^3 + y^3 = Cx^2$ 49. $2x^3 - 6 \ln x + 3e^{y^2} - 5 = 0$
 29. $y = Cx$ 51. $\cos y = \sec x$
 31. $x^3 e^{y^3/x^3} = C$ 53. $y = (ax^2 + 2abx + y_0^2)^{1/2}$
 33. $e^y - e^x = C$ 55. $R = (10x^2 - x^3)^{1/3}$
 35. $x = Ce^{y/x}$ 57. $y = \sqrt{8x^4 + x^2}$
 37. $y = 3x$ 59. a. \$7460
 b. \$5421.69
 39. $y^2 = x^2 - x$ 61. 329.8
 41. $\ln x + e^{-y/x} = 1$ 63. $y = 225 + a - ax - k \ln x$
 43. $x^2 - 4y^2 = 9$
 45. $2\sqrt{\frac{x}{y}} - \ln y = 4$

EXACT DIFFERENTIAL EQUATIONS

The total derivative of a function of x and y , say $F(x, y)$, is given by

$$dF(x, y) = \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy$$

Thus the differential equation

$$\frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy = 0$$

has the general solution $F(x, y) = C$ and is said to be an *exact differential equation*.

A differential equation of the general form

$$M(x, y) dx + N(x, y) dy = 0$$

is said to be an exact differential equation if $M(x, y) dx + N(x, y) dy$ is the total derivative of some function $F(x, y)$; $M(x, y)$ and $N(x, y)$ are then the partial derivatives of $F(x, y)$ with respect to x and y , respectively.

If the second-order mixed partial derivatives of $F(x, y)$ exist and are continuous, then

$$\frac{\partial}{\partial y} \left(\frac{\partial F}{\partial x} \right) = \frac{\partial}{\partial x} \left(\frac{\partial F}{\partial y} \right)$$

Thus if a differential equation of the form $M(x, y) dx + N(x, y) dy = 0$ is exact,

$$\frac{\partial}{\partial y} M(x, y) = \frac{\partial}{\partial x} N(x, y)$$

It can be shown that this is also a sufficient condition for exactness; that is,

$$\frac{\partial}{\partial y} M(x, y) = \frac{\partial}{\partial x} N(x, y) \Leftrightarrow M(x, y) dx + N(x, y) dy = 0 \text{ is exact}$$

If a differential equation is exact, its solution can be found by the following method.

1. Integrate $M(x, y)$ with respect to x replacing the usual constant of integration with a function $f(y)$ of y .

$$F(x, y) = \int M(x, y) dx = G(x, y) + f(y)$$

2. Differentiate $F(x, y) = G(x, y) + f(y)$ obtained from step 1 with respect to y and compare this with $N(x, y)$ from the differential equation to be solved to obtain the value of $\frac{\partial}{\partial y} f(y)$.

$$\frac{\partial G}{\partial y} + \frac{\partial}{\partial y} f(y) = \frac{\partial N}{\partial y}$$

$$\frac{\partial}{\partial y} f(y) = \frac{\partial N}{\partial y} - \frac{\partial G}{\partial y}$$

3. Integrate $\frac{\partial}{\partial y} f(y)$ with respect to y to obtain $f(y)$.

$$\int \frac{\partial}{\partial y} f(y) dy = f(y)$$

It is not necessary to include the usual constant of integration, since it is introduced in the final step of the solution.

4. The solution, from steps 1 and 3, is

$$F(x, y) = G(x, y) + f(y) + C = 0$$

clearly the solution may also be obtained by integrating first with respect to y .

EXAMPLE

Solve the equation $[y(2 + x^3y)/x^3] dx = [(1 - 2x^3y)/x^2] dy$.

$$\frac{\partial}{\partial y} \left(\frac{y(2 + x^3y)}{x^3} \right) = \frac{2 + 2x^3y}{x^3} = \frac{\partial}{\partial x} \left(-\frac{1 - 2x^3y}{x^2} \right), \text{ so the equation is exact.}$$

$$\frac{\partial F}{\partial x} = \frac{y(2 + x^3y)}{x^3}$$

$$F(x, y) = -\frac{y}{x^2} + y^2x + f(y) \quad (\text{step 1})$$

$$\frac{\partial F}{\partial y} = -\frac{1}{x^2} + 2xy + \frac{\partial}{\partial y} f(y) = -\frac{1}{x^2} + 2xy$$

so

$$\frac{\partial}{\partial y} f(y) = 0 \quad (\text{step 2})$$

$$f(y) = C \quad (\text{step 3})$$

and

$$F(x, y) = -\frac{y}{x^2} + y^2x + C = 0 \quad (\text{step 4})$$

Thus

$$x^3 y^2 - y = Cx^2$$

is the solution.

EXAMPLE

Solve the equation $(x^2 - x + y^2) dx - (ye^y - 2xy) dy = 0$.

$$\frac{\partial}{\partial y} (x^2 - x + y^2) = 2y = \frac{\partial}{\partial x} (-ye^y + 2xy), \text{ so the equation is exact.}$$

$$\frac{\partial F}{\partial x} = x^2 - x + y^2$$

$$F(x, y) = \frac{x^3}{3} - \frac{x^2}{2} + xy^2 + f(y) \quad (\text{step 1})$$

$$\frac{\partial F}{\partial y} = 2xy + \frac{\partial}{\partial y} f(y) = -ye^y + 2xy$$

so

$$\frac{\partial}{\partial y} f(y) = -ye^y \quad (\text{step 2})$$

$$f(y) = -(y - 1)e^y \quad (\text{step 3})$$

and

$$F(x, y) = \frac{x^3}{3} - \frac{x^2}{2} + xy^2 - (y - 1)e^y + C = 0 \quad (\text{step 4})$$

Thus

$$2x^3 - 3x^2 + 6xy^2 - 6(y - 1)e^y = C$$

is the solution.

NOTE: Alternatively, the terms of the original equation can be considered in two sets: $x^2 dx - x dx - ye^y dy$ is exact by inspection and can be integrated by the usual methods; $y^2 dx + 2xy dy$ can be shown to be exact and can be integrated by the method above. When this method is used, $f(y)$ is a constant, as illustrated in the following example.

EXAMPLE

Solve the equation $[(2xy - 1)/y] dx + [(x + 3y)/y^2] dy = 0$ and find the particular solution if $y = 1$ when $x = 2$.

$$\underbrace{2x dx + \frac{3}{y} dy}_{\text{exact, integrable by usual methods to obtain } x^2 + 3 \ln y} - \underbrace{\frac{1}{y} dx + \frac{x}{y^2} dy}_{\frac{\partial}{\partial y} \left(-\frac{1}{y} \right) = \frac{1}{y^2} = \frac{\partial}{\partial x} \left(\frac{x}{y^2} \right)} = 0$$

$$\frac{\partial}{\partial y} \left(-\frac{1}{y} \right) = \frac{1}{y^2} = \frac{\partial}{\partial x} \left(\frac{x}{y^2} \right)$$

so the equation is exact.

$$\frac{\partial F}{\partial x} = -\frac{1}{y}$$

$$F(x, y) = -\frac{x}{y} + f(y) \quad (\text{step 1})$$

$$\frac{\partial F}{\partial y} = \frac{x}{y^2} + \frac{\partial}{\partial y} f(y) = \frac{x}{y^2}$$

so

$$\frac{\partial}{\partial y} f(y) = 0 \quad (\text{step 2})$$

$$f(y) = C \quad (\text{step 3})$$

and

$$F(x, y) = -\frac{x}{y} + C \quad (\text{step 4})$$

Thus

$$x^2 + 3 \ln y - \frac{x}{y} = C$$

and

$$x^2 y - x + 3y \ln y = Cy$$

is the solution.

If $y = 1$ when $x = 2$,

$$4 - 2 = C$$

$$C = 2$$

and

$$x^2 y - x - 2y + 3y \ln y = 0$$

is the particular solution for $x = 2, y = 1$.

EXAMPLE

The change in price, y , with change in quantity demanded, x , of a particular commodity is given by

$$\frac{dy}{dx} = -\frac{2xy + 24x}{x^2 + 16}$$

Find the relationship between price and quantity demanded if the price is 7.5 when the quantity demanded is 4.

$$(2xy + 24x) dx + (x^2 + 16) dy = 0$$

$$\frac{\partial}{\partial y} (2xy + 24x) = 2x = \frac{\partial}{\partial x} (x^2 + 16)$$

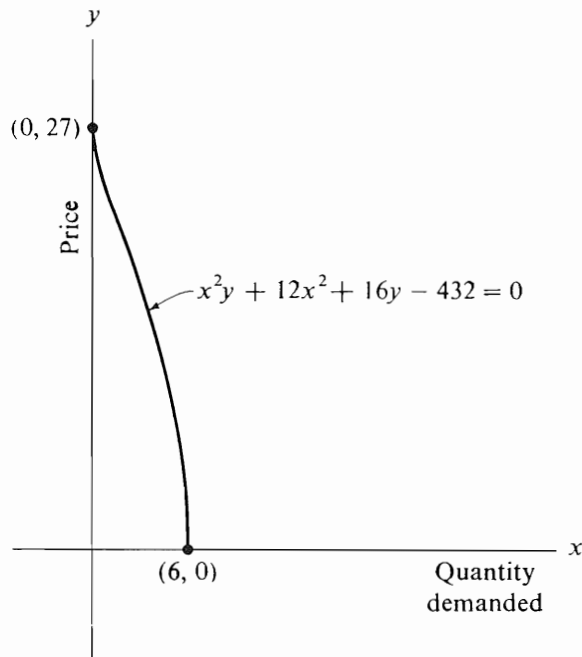
so the equation is exact.

$$\frac{\partial F}{\partial x} = 2xy + 24x$$

$$F(x, y) = x^2 y + 12x^2 + f(y) \quad (\text{step 1})$$

$$\frac{\partial F}{\partial y} = x^2 + \frac{\partial}{\partial y} f(y) = x^2 + 16$$

Figure 5.4



so

$$\frac{\partial}{\partial y} f(y) = 16 \quad (\text{step 2})$$

$$f(y) = 16y \quad (\text{step 3})$$

and

$$F(x, y) = x^2y + 12x^2 + 16y + C = 0 \quad (\text{step 4})$$

thus

$$x^2y + 12x^2 + 16y = C$$

is the solution.

If $y = 7.5$ when $x = 4$,

$$120 + 192 + 120 = C$$

$$C = 432$$

and

$$x^2y + 12x^2 + 16y - 432 = 0$$

is the particular solution for $x = 4$, $y = 7.5$. (See Figure 5.4.)

LINEAR DIFFERENTIAL EQUATIONS

In some cases a differential equation which is not exact can be made exact by multiplication by a factor; such a factor is called an *integrating factor*, since it permits the equation to be integrated. In general, determination of the appropriate integrating factor for a given differential equation may not be an easy matter; however, it can

be shown that any linear first-order differential equation has the integrating factor $e^{\int P(x) dx}$.

An equation which can be written in the form

$$\frac{dy}{dx} + yP(x) = Q(x)$$

is a *linear differential equation of first order*. If this equation is multiplied by the integrating factor $e^{\int P(x) dx}$, the resulting equation

$$e^{\int P(x) dx} \frac{dy}{dx} + yP(x)e^{\int P(x) dx} = Q(x)e^{\int P(x) dx}$$

can be integrated with respect to x to obtain

$$ye^{\int P(x) dx} = \int e^{\int P(x) dx} Q(x) dx + C$$

since

$$\frac{d}{dx} (ye^{\int P(x) dx}) = e^{\int P(x) dx} \frac{dy}{dx} + yP(x)e^{\int P(x) dx}$$

Thus the equation

$$\frac{dy}{dx} + yP(x) = Q(x)$$

has as its solution

$$y = e^{-\int P(x) dx} \left[\int e^{\int P(x) dx} Q(x) dx + C \right]$$

Similarly, the equation

$$\frac{dx}{dy} + xP(y) = Q(y)$$

has $e^{\int P(y) dy}$ as an integrating factor and its solution is

$$x = e^{-\int P(y) dy} \left[\int e^{\int P(y) dy} Q(y) dy + C \right]$$

NOTE: Every differential equation of first order and first degree, which has a solution, can be shown to have an infinite number of integrating factors. However, determining these integrating factors may not be easy; in general, only linear first-order differential equations are solved using integrating factors. As discussed in the following section, first-order differential equations which are linear in a function of a variable can also be integrated using integrating factors.

EXAMPLE

Solve the equation $x^2 dy + (y - 2xy - 2x^2) dx = 0$.

Writing the equation in the form $\frac{dy}{dx} + yP(x) = Q(x)$,

$$\frac{dy}{dx} + y\left(\frac{1-2x}{x^2}\right) = 2$$

The integrating factor is

$$e^{\int [(1-2x)/x^2] dx} = e^{-(1/x) - \ln x^2} = \frac{e^{-1/x}}{x^2}$$

and

$$\begin{aligned} \frac{ye^{-1/x}}{x^2} &= \int \frac{2e^{-1/x}}{x^2} dx \\ &= 2e^{-1/x} + C \\ y &= 2x^2 + Cx^2e^{1/x} \\ &= x^2(2 + Ce^{1/x}) \end{aligned}$$

EXAMPLE

Solve the equation $y \frac{dx}{dy} = 2ye^{3y} + x(3y + 2)$.

Writing the equation in the form $\frac{dx}{dy} + xP(y) = Q(y)$,

$$\frac{dx}{dy} - x\left(\frac{3y+2}{y}\right) = 2e^{3y}$$

The integrating factor is

$$e^{\int -[(3y+2)/y] dy} = e^{-3y-2 \ln y} = \frac{e^{-3y}}{y^2}$$

and

$$\begin{aligned} \frac{xe^{-3y}}{y^2} &= \int \frac{2}{y^2} dy \\ &= -\frac{2}{y} + C \\ x &= -2ye^{3y} + Cy^2e^{3y} \\ &= ye^{3y}(Cy - 2) \end{aligned}$$

EXAMPLE

Solve the equation $2y dx = (y^4 + x) dy$ and find the particular solution if $y = 1$ when $x = 0$.

Writing the equation in the form $\frac{dx}{dy} + xP(y) = Q(y)$,

$$\frac{dx}{dy} - x\left(\frac{1}{2y}\right) = \frac{y^3}{2}$$

The integrating factor is

$$e^{\int -dy/2y} = e^{-(1/2) \ln y} = y^{-1/2}$$

and

$$xy^{-1/2} = \frac{1}{2} \int y^{5/2} dy$$

$$= \frac{1}{7} y^{7/2} + C$$

$$x = \frac{1}{7} y^4 + C y^{1/2}$$

$$7x = y^4 + C y^{1/2}$$

$$\text{If } y = 1 \text{ when } x = 0,$$

$$0 = 1 + C$$

$$C = -1$$

and

$$7x = y^4 - y^{1/2}$$

is the particular solution for $x = 0, y = 1$.

EXAMPLE

A manufacturing company has found that the cost, c , of operating and maintaining its equipment is related to the length, x , of the interval between overhauls by the equation

$$\frac{dc}{dx} - \frac{b-1}{x} c = -\frac{ba}{x^2}$$

where a and b are constants. Find c as a function of x if $c = c_0$ when $x = x_0$.

Writing the equation in the form $\frac{dc}{dx} + cP(X) = Q(x)$,

$$\frac{dc}{dx} - c \left(\frac{(b-1)}{x} \right) = -\frac{ba}{x^2}$$

The integrating factor is

$$e^{\int -[(b-1)/x] dx} = e^{-(b-1) \ln x} = x^{-(b-1)}$$

and

$$cx^{-(b-1)} = -ba \int x^{-b-1}$$

$$= ax^{-b} + C$$

$$c = \frac{a}{x} + Cx^{b-1}$$

If $c = c_0$ when $x = x_0$,

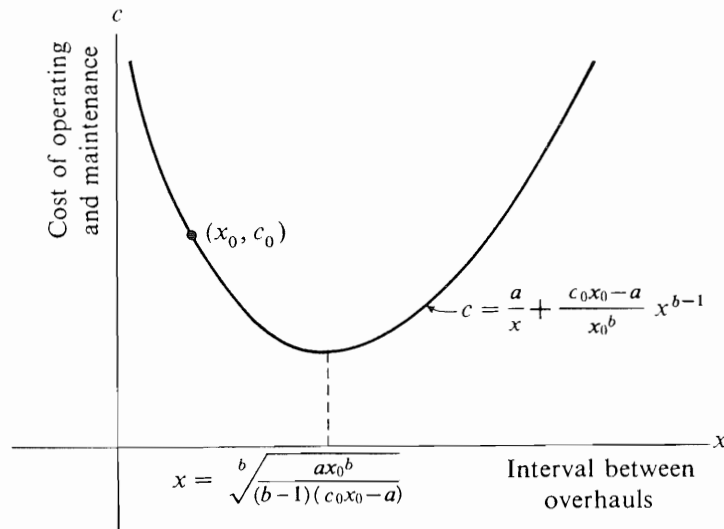
$$c_0 = \frac{a}{x_0} + Cx_0^{b-1}$$

$$C = \frac{c_0 x_0 - a}{x_0^b}$$

$$c = \frac{a}{x} + \frac{C_0 x_0 - a}{x_0^b} x^{b-1}$$

(See Figure 5.5.)

Figure 5.5



DIFFERENTIAL EQUATIONS LINEAR IN A FUNCTION OF y OR IN A FUNCTION OF x

An equation which can be written in the form

$$\frac{d}{dx} f(y) + f(y)P(x) = Q(x)$$

is a linear differential equation in the variable $f(y)$; that is, it is of the first degree in $f(y)$ and $\frac{d}{dy} f(y)$. Its solution is thus given by

$$f(y) = e^{-\int P(x) dx} \left[\int e^{\int P(x) dx} Q(x) dx + C \right]$$

Similarly, an equation which can be written in the form

$$\frac{d}{dy} f(x) + f(x)P(y) = Q(y)$$

is a linear differential equation in the variable $f(x)$; that is, it is of the first degree in $f(x)$ and $\frac{d}{dx} f(x)$. Its solution is thus given by

$$f(x) = e^{-\int P(y) dy} \left[\int e^{\int P(y) dy} Q(y) dy + C \right]$$

EXAMPLE

Solve the equation $y dx + x(1 - x^2 y^4) dy = 0$.

Writing the equation in the form $\frac{d}{dy} f(x) + f(x)P(y) = Q(y)$,

$$\frac{dx}{dy} + \frac{x}{y} = x^3 y^3$$

$$x^{-3} \frac{dx}{dy} + x^{-2} \left(\frac{1}{y} \right) = y^3$$

$$-2x^{-3} \frac{dx}{dy} + x^{-2} \left(-\frac{2}{y} \right) = -2y^3$$

$$\frac{d}{dy} (x^{-2}) + x^{-2} \left(-\frac{2}{y} \right) = -2y^3$$

The integrating factor is

$$e^{\int -(2/y) dy} = e^{-2 \ln y} = y^{-2}$$

and

$$x^{-2} = y^2 \left[-2 \int y^{-2} y^3 dy + C \right]$$

$$= y^2 \left[-2 \int y dy + C \right]$$

$$= y^2 [-y^2 + C]$$

$$1 = -x^2 y^4 + C x^2 y^2$$

$$1 + x^2 y^4 = C x^2 y^2$$

EXAMPLE

Solve the equation $\frac{dy}{dx} + xy \ln y = xye^{-x^2}$ and find the particular solution if $y = 1$ when $x = 0$.

Writing the equation in the form $\frac{d}{dx} f(y) + f(y)P(x) = Q(x)$,

$$\frac{1}{y} \frac{dy}{dx} + (\ln y)(x) = xe^{-x^2}$$

$$\frac{d}{dx} (\ln y) + (\ln y)(x) = xe^{-x^2}$$

The integrating factor is

$$e^{\int x dx} = e^{x^2/2}$$

and

$$\ln y = e^{-x^2/2} \left[\int x e^{x^2/2} e^{-x^2} dx + C \right]$$

$$= e^{-x^2/2} \left[\int x e^{-x^2/2} dx + C \right]$$

$$= e^{-x^2/2} [-e^{-x^2/2} + C]$$

$$= -e^{-x^2} + C e^{-x^2/2}$$

If $y = 1$ when $x = 0$,

$$0 = -1 + C$$

$$C = 1$$

and

$$\begin{aligned}\ln y &= -e^{-x^2} + e^{-x^2/2} \\ &= e^{-x^2/2}(1 - e^{-x^2/2})\end{aligned}$$

is the particular solution for $x = 0, y = 1$.

EXAMPLE

The cost control unit of a large public accounting firm has found, as the company has increased in size, that the average monthly cost, y , of office supplies is related to the number, x , of employees (in addition to the unit director) by the equation

$$\frac{dy}{dx} + 2y = y^2 e^{-x}$$

Find y as a function of x if $y = 3$ when $x = 0$.

Writing the equation in the form $\frac{d}{dx} f(y) + f(y)P(x) = Q(x)$,

$$y^{-2} \frac{dy}{dx} + 2y^{-1} = e^{-x}$$

$$-y^{-2} \frac{dy}{dx} - 2y^{-1} = -e^{-x}$$

$$\frac{d}{dx}(y^{-1}) - 2y^{-1} = -e^{-x}$$

The integrating factor is

$$e^{\int -2 dx} = e^{-2x}$$

and

$$\begin{aligned}y^{-1} &= e^{2x} \left[\int -e^{-2x} e^{-x} dx + C \right] \\ &= e^{2x} \left[-\int e^{-3x} dx + C \right] \\ &= e^{2x} \left[\frac{e^{-3x}}{3} + C \right] \\ &= \frac{e^{-x}}{3} + C\end{aligned}$$

$$3e^x = y + Cy e^x$$

If $y = 3$ when $x = 0$,

$$3 = 3 + 3C$$

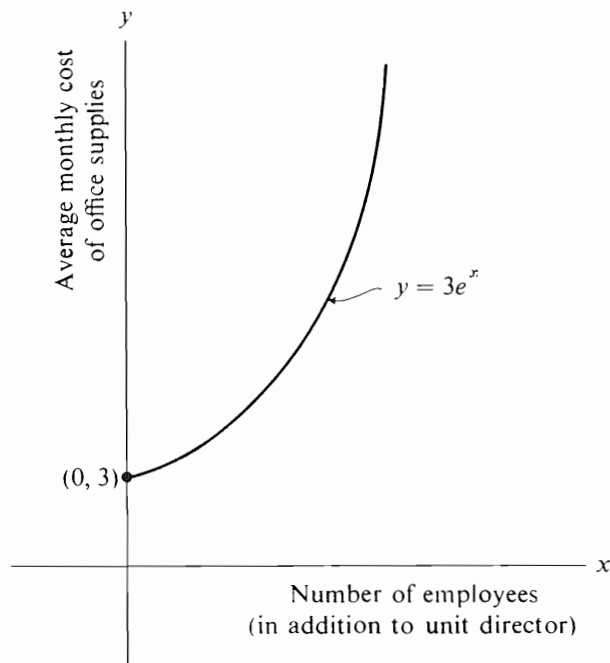
$$C = 0$$

and

$$y = 3e^x$$

is the particular solution for $x = 0, y = 3$. (See Figure 5.6.)

Figure 5.6



NOTE: In general, there is no rule for determining the function $f(y)$ or $f(x)$ in terms of which a differential equation is, or can be made, linear. The equation must be put in the form $g(y) \frac{dy}{dx} + f(y)P(x) = Q(x)$ or the form $g(x) \frac{dx}{dy} + f(x)P(y) = Q(y)$ so that it can be determined by inspection whether the differential equation can be made linear in $f(y)$ or $f(x)$. However, there is one class of differential equations for which the appropriate function is known. An equation of the form

$$\frac{dy}{dx} + yP(x) = y^n Q(x)$$

or

$$\frac{dx}{dy} + xP(y) = x^n Q(y)$$

can be made linear in y^{n-1} or x^{n-1} by multiplication by $(1-n)y^{-n}$ or $(1-n)x^{-n}$, respectively. Equations of this type are known as *Bernoulli equations*. The first and last examples of this section involve Bernoulli equations.

5.5 Differential Equations of Higher Order and/or Higher Degree

Except for special cases, methods of solution of differential equations of the second order or degree are somewhat complicated and, in general, solution becomes more difficult as the order or degree of the differential equation increases. In many cases, methods for obtaining an exact solution are not known and numerical methods for

obtaining approximate solutions must be used. It should be noted that there are, in fact, some types of differential equations of first order and first degree for which exact solutions are not known. With electronic computers, highly accurate approximate solutions can be obtained; these solutions may sometimes be preferred, for reasons of convenience, even though methods for exact solution are known.

PROBLEMS

Find the general solution for each of the following differential equations.

1. $x \, dy - 3y \, dx = x^2 \, dx$

2. $\frac{dy}{dx} + 2xy - 2xe^{-x^2} = 0$

3. $y \frac{dx}{dy} + (1 + y)x - e^y = 0$

4. $y \frac{dx}{dy} + 2x - 3y = 0$

5. $x(6xy + 5) \, dx + (2x^3 + 3y) \, dy = 0$

6. $(a_1x + b_1y + c_1) \, dx + (b_1x + b_2y + c_2) \, dy = 0$

7. $y \, dx + 3x^2e^x \, dx = dy$

8. $(ye^{xy} + 2xy) \, dx + (xe^{xy} + x^2) \, dy = 0$

9. $(x + 1) \frac{dy}{dx} + 2y = e^x(x + 1)^{-1}$

10. $(y \cos x - 2 \sin y) \, dx = (2x \cos y - \sin x) \, dy$

11. $(3x^2y + xy^2 + e^x) \, dx + (x^3 + x^2y + \sin y) \, dy = 0$

12. $2y \sin xy \, dx + (2x \sin xy + y^3) \, dy = 0$

13. $(3y \sin x - \cos y) \, dx + (x \sin y - 3 \cos x) \, dy = 0$

14. $(y^2 \csc^2 x + 6xy - 2) \, dx = (2y \cot x - 3x^2) \, dy$

15. $\cos y \, dx - (x \sin y - y^2) \, dy = 0$

16. $2 \left(\frac{y}{x^3} + \frac{x}{y^2} \right) \, dx = \left(\frac{1}{x^2} + \frac{2x^2}{y^3} \right) \, dy$

17. $\frac{dy}{dx} + 2y - x = 0$

23. $y^3 \frac{dy}{dx} + xy^4 = xe^{-x^2}$

18. $x \frac{dy}{dx} = x^3 + y$

24. $xy \, dy = (x^2 - y^2) \, dx$

25. $(x + y) \, dx + (x - 2y) \, dy = 0$

19. $3 \frac{dy}{dx} + \frac{2y}{x+1} = \frac{x}{y^2}$

26. $\frac{dy}{dx} + \left(y - \frac{1}{y} \right) x = 0$

20. $\frac{dx}{dy} + x = e^{-y}$

27. $\frac{ds}{dt} - s \cot t = 1 - (t + 2) \cot t$

21. $\frac{dy}{dx} - xy = y^{1/2}xe^{x^2}$

28. $\frac{ds}{dt} + \frac{s}{t} = \cos t + \frac{\sin t}{t}$

22. $\frac{dy}{dx} - xy = \frac{x}{y}$

29. $\frac{dy}{dx} + y = \cos x - \sin x$

30. $(2x \cos y - e^x) dx - x^2 \sin y dy = 0$
31. $(y \sin x + xy \cos x) dx + (\sin y + e^y) dy = 0$
32. $\frac{dy}{dx} - xy = e^{(1/2)x^2} \cos x$ 33. $\cos \theta \frac{dr}{d\theta} = 2 + 2r \sin \theta$
34. $\cos x \sec y dx + \sin x \sin y \sec^2 y dy = 0$
35. $\frac{ds}{dt} + s \tan t = 2t + t^2 \tan t$
36. $\cos y dy + (\sin y - 1) \cos x dx = 0$
37. $x \tan^2 y dy + x dy = (2x^2 + \tan y) dx$
38. $3y^2 \frac{dy}{dx} - xy^3 = e^{(1/2)x^2} \cos x$
39. $(\ln y + y^3 + ye^{xy}) dx + \left(\frac{x}{y} + 3xy^2 + xe^{xy}\right) dy = 0$
40. $dx + (2xy - 4xy^3 - 2y^3e^{y^4}) dy = 0$
41. $\left(y^2 + \frac{1}{y} + ye^y\right) dx + \left(2xy - \frac{x}{y^2} + xe^y + xye^y\right) dy = 0$
42. $y^2 \frac{dx}{dy} + (y^2 + 2y)x - 1 = 0$ 47. $x^2 \frac{dy}{dx} + y^2 = xy$
43. $x dy = (5y + x + 1) dx$ 48. $\sin \theta d\theta + \cos \theta dt = te^t dt$
44. $\frac{xy + 1}{y} dx + \frac{2y - x}{y^2} dy = 0$ 49. $\csc y \cot y dy = (\csc y + e^x) dx$
45. $\frac{dx}{dy} = x + e^y$ 50. $x \frac{dy}{dx} + y = y^2 x^2 \cos x$
46. $x \frac{dy}{dx} + 2y = 3x^3 y^{4/3}$ 51. $\frac{dx}{dy} - x \cot y = e^y(1 - \cot y)$

Find the particular solution for each of the following differential equations under the given conditions.

52. $\frac{dy}{dx} + y \tan x = \sec x$, $y = -1$ when $x = 0$
53. $dy + (y \cot x - \sec x) dx = 0$, $y = 1$ when $x = 0$
54. $(y^2 + 1) \frac{dx}{dy} + 2xy = y^2$, $y = -1$ when $x = 0$
55. $dy = x(1 - e^{2y-x^2}) dx$, $y = 0$ when $x = 0$
56. $\left(x^2 + \frac{1}{x} + ye^{xy}\right) dx + (e^y + 3y^2 + xe^{xy}) dy = 0$, $y = 0$ when $x = 1$
57. $(1 - x^2) \frac{dy}{dx} + xy = x(1 - x^2)y^{1/2}$, $y = 1$ when $x = 0$
58. $\frac{dy}{dx} + y = y^2 e^{-x}$, $y = 2$ when $x = 0$
59. $y dx + 2(x - 2y^2) dy = 0$, $y = -1$ when $x = 2$
60. $(y + y^3) dx + 4(xy^2 - 1) dy = 0$, $y = 1$ when $x = 0$

61. $2y \, dx = (x^2 y^4 + x) \, dy$, $y = 1$ when $x = 1$
62. $(x^2 - 1) \frac{dy}{dx} + (x^2 - 1)^2 + 4y = 0$, $y = -6$ when $x = 0$
63. $(ye^x - 2x) \, dx + e^x \, dy = 0$, $y = 6$ when $x = 0$
64. $\left(\frac{2}{y} - \frac{y}{x^2}\right) \, dx + \left(\frac{1}{x} - \frac{2x}{y^2}\right) \, dy = 0$, $y = 2$ when $x = 1$
65. $(2y - xy - 3) \, dx + x \, dy = 0$, $y = 1$ when $x = 1$
66. $(ye^y + 2x + y) \, dx + (xe^y + xye^y + 3y^2 + x) \, dy = 0$, $y = 0$ when $x = 2$
67. $\left(3x^2 y - \frac{2y}{x} - 6x^4\right) \, dx + dy = 0$, $y = 0$ when $x = -1$
68. $dr = (1 + 2r \cot \theta) \, d\theta$, $r = 3$ when $\theta = \pi/2$
69. The change in net profit, P , as advertising expenditure, x , changes is given by the equation

$$\frac{dP}{dx} = k - a(P + x)$$

where a and k are constants. Find P as a function of x if $P = P_0$ when $x = 0$.

70. The change in the cost of ordering and holding, C , as quantity, x , changes is given by the equation

$$\frac{dC}{dx} = a - \frac{C}{x}$$

where a is a constant. Find C as a function of x if $C = C_0$ when $x = x_0$.

71. Manufacturing and marketing costs, C , are related to the number of items, x , by the equation

$$\frac{dC}{dx} + aC = b + kx$$

where a , b , and k are constants. Find C as a function of x if $C = 0$ when $x = 0$.

72. The change in consumption, c , of a particular commodity as income, I , changes is given by the equation

$$\frac{dc}{dI} = c + ke^I$$

where k is a constant. Find c as a function of I if $c = c_0$ when $I = 0$.

Answers to Odd-Numbered Problems

- | | |
|---|---|
| 1. $y = -x^2 + Cx^3$ | 21. $3y^{1/2} = e^{x^2} + Ce^{x^2/4}$ |
| 3. $2xye^y - e^{2y} = C$ | 23. $e^{2x^2}y^4 = 2e^{x^2} + C$ |
| 5. $4x^3y + 5x^2 + 3y^2 = C$ | 25. $x^2 + 2xy - 2y^2 = C$ |
| 7. $y = x^3e^x + Ce^x$ | 27. $s = t + 2 + C \sin t$ |
| 9. $y(x+1)^2 = e^x + C$ | 29. $y = \cos x + Ce^{-x}$ |
| 11. $2x^3y + x^2y^2 + 2e^x - 2 \cos y = C$ | 31. $xy \sin x - \cos y + e^y = C$ |
| 13. $3y \cos x + x \cos y = C$ | 33. $r \cos^2 \theta = 2 \sin \theta + C$ |
| 15. $y^3 + 3x \cos y = C$ | 35. $s = t^2 + C \cos t$ |
| 17. $y = \frac{1}{2}x - \frac{1}{4} + Ce^{-2x}$ | 37. $\tan y = 2x^2 + Cx$ |
| 19. $12(x+1)^2y^3 = 3x^4 + 8x^3 + 6x^2 + C$ | 39. $x \ln y + xy^3 + e^{xy} = C$ |

41. $x\left(y^2 + \frac{1}{y} + ye^y\right) = C$

43. $20y + 5x + 4 = Cx^5$

45. $x = ye^y + Ce^y$

47. $x = y \ln Cx$

49. $2e^x \csc y + e^{2x} = C$

51. $x = e^y + C \sin y$

53. $y \sin x + \ln \cos x = 0$

55. $x^2 + 1 = e^{x^2 - 2y}$

57. $3y^{1/2} + 1 - x^2 = 4(1 - x^2)^{1/4}$

59. $y^2(x - y^2) = 1$

61. $10x = (9 + xy^4)y^{1/2}$

63. $ye^x - x^2 = 6$

65. $x^2y + 3(x + 1) = 7e^{x-1}$

67. $y = 2x^2 - 2x^2e^{-x^3-1}$

69. $P = \frac{k+1}{a} - x + \left(P_0 - \frac{k+1}{a}\right)e^{-ax}$

71. $C = \frac{ab-k}{a^2}(1 - e^{-ax}) + \frac{k}{a}x$

5.6 Applications of Differential Equations in Economic Models

There are two general types of economic models, *static* and *dynamic*. Static models concern equilibrium situations, that is, situations which if attained will be maintained. Dynamic models concern situations which change over time. In dynamic models, time enters either explicitly as a variable or implicitly in the form of lagged variables. The models discussed in this chapter are very simple dynamic economic models expressed in terms of differential equations.

There are two general classes of variables in economic models, referred to as *endogenous* and *exogenous* variables. Endogenous variables are those variables whose values or levels are to be predicted or explained; exogenous variables are assumed to be determined and known in advance and can be regarded as constant in the model. Endogenous and exogenous come from Greek words meaning "generated from inside" and "generated from outside," respectively. Endogenous variables are predicted from the model; exogenous variables are determined outside the model.

Usually a model is first written in terms of *structural equations* which express relationships among endogenous and exogenous variables. This set of structural equations is then solved, if possible, for what are called *reduced form* equations, each of which expresses an endogenous variable as a function of exogenous variables and parameters. A model is solved by obtaining a reduced form equation for each endogenous variable in the model. Solution of economic models is illustrated in this section for the Domar macro model, two Domar debt models, the Evans price adjustment model, and an income-consumption-investment model.

DOMAR MACRO MODEL

The following very simple macro model was proposed by E. D. Domar.

$$S(t) = \alpha y(t)$$

$$I(t) = \beta \frac{dy}{dt}$$

$$S(t) = I(t)$$

$$y(0) = y_0$$

$$\alpha > 0 \quad \beta > 0$$

where S is saving, I is investment, y is income, and each of these endogenous variables is a function of time.

The first equation states that saving is a fixed proportion of income; the second equation states that investment is proportional to the rate of change of income over time; the third equation states that saving equals investment; the fourth equation states the initial condition. From these relationships, specific functions expressing variations in the variables over time can be obtained.

Since $S(t) = I(t)$, the differential equation

$$\frac{dy}{dt} - \frac{\alpha}{\beta}y = 0$$

is obtained for solution.

$$\frac{1}{y} \frac{dy}{dt} = \frac{\alpha}{\beta}$$

$$\ln y = \frac{\alpha}{\beta}t + C$$

$$y = Ce^{(\alpha/\beta)t}$$

If $y = y_0$ at $t = 0$, then

$$y_0 = C$$

and the particular solution is

$$y = y_0 e^{(\alpha/\beta)t}$$

Note that this solution gives income, y , as a function of time, t . Since $\alpha > 0$, $\beta > 0$, the function has an increasing positive slope, the rate of increase depending on α/β . Solutions for the remaining variables of the model, I and S , are as follows:

$$I = S = \alpha y = \alpha y_0 e^{(\alpha/\beta)t}$$

DOMAR DEBT MODEL

Domar uses a set of models similar to the macro model above to express relationships between national income and national debt. Consider the model

$$\frac{dD}{dt} = \alpha y(t)$$

$$\frac{dy}{dt} = \beta$$

$$y(0) = y_0$$

$$D(0) = D_0$$

$$\alpha > 0 \quad \beta > 0$$

where D is national debt and y is national income (both endogenous). In this model, national income increases at a constant rate β over time and the rate of increase of national debt is a fixed proportion of national income. The third and fourth equations state the initial conditions. Integrating the second equation,

$$y = \beta t + C$$

Since $y = y_0$ when $t = 0$, $C = y_0$ and

$$y = \beta t + y_0$$

Substituting in the first equation of the model,

$$\frac{dD}{dt} = \alpha\beta t + \alpha y_0$$

$$D = \frac{1}{2}\alpha\beta t^2 + \alpha y_0 t + C$$

Since $D = D_0$ when $t = 0$, $C = D_0$ and

$$D = \frac{1}{2}\alpha\beta t^2 + \alpha y_0 t + D_0$$

The solution of the model is thus

$$D(t) = \frac{1}{2}\alpha\beta t^2 + \alpha y_0 t + D_0$$

$$y(t) = \beta t + y_0$$

Domar was interested in the ratio of national debt to national income

$$\frac{D(t)}{y(t)} = \frac{\frac{1}{2}\alpha\beta t^2 + \alpha y_0 t + D_0}{\beta t + y_0}$$

or

$$\frac{D(t)}{y(t)} = \frac{D_0}{\beta t + y_0} + \frac{\alpha y_0 t}{\beta t + y_0} + \frac{\frac{1}{2}\alpha\beta t^2}{\beta t + y_0}$$

As $t \rightarrow \infty$,

$$\frac{D_0}{\beta t + y_0} \rightarrow 0$$

$$\frac{\alpha y_0 t}{\beta t + y_0} \rightarrow \frac{\alpha y_0}{\beta} \quad (\text{a constant})$$

$$\frac{\frac{1}{2}\alpha\beta t^2}{\beta t + y_0} \rightarrow \infty$$

Thus, as $t \rightarrow \infty$, $D(t)/y(t) \rightarrow \infty$ and, for this model, the ratio of national debt to national income increases over time without limit.

A SECOND DOMAR DEBT MODEL

A second version of the Domar debt model is obtained by modifying the second equation in the above model, so that income increases by a constant proportion of income, and leaving the other equations unchanged.

$$\frac{dD}{dt} = \alpha y(t)$$

$$\frac{dy}{dt} = \beta y(t)$$

$$y(0) = y_0$$

$$D(0) = D_0$$

$$\alpha > 0 \quad \beta > 0$$

Integrating the second equation,

$$\frac{1}{y} \frac{dy}{dt} = \beta$$

$$\ln y = \beta t + C$$

$$y = C e^{\beta t}$$

Since $y = y_0$ when $t = 0$, $C = y_0$ and

$$y = y_0 e^{\beta t}$$

Substituting in the first equation of the model,

$$\frac{dD}{dt} = \alpha y_0 e^{\beta t}$$

$$D = \frac{\alpha}{\beta} y_0 e^{\beta t} + C$$

Since $D = D_0$ when $t = 0$, $C = D_0 - \frac{\alpha}{\beta} y_0$ and

$$D = D_0 - \frac{\alpha}{\beta} y_0 + \frac{\alpha}{\beta} y_0 e^{\beta t}$$

or

$$D = D_0 + \frac{\alpha}{\beta} y_0 (e^{\beta t} - 1)$$

The solution of the model is thus

$$D(t) = D_0 + \frac{\alpha}{\beta} y_0 (e^{\beta t} - 1)$$

$$y(t) = y_0 e^{\beta t}$$

The ratio of national debt to national income is

$$\frac{D(t)}{y(t)} = \frac{D_0}{y_0 e^{\beta t}} + \frac{\alpha}{\beta} \left(1 - \frac{1}{e^{\beta t}} \right)$$

As $t \rightarrow \infty$,

$$\frac{D_0}{y_0 e^{\beta t}} \rightarrow 0$$

$$\frac{\alpha}{\beta} \left(1 - \frac{1}{e^{\beta t}} \right) \rightarrow \frac{\alpha}{\beta}$$

Thus, as $t \rightarrow \infty$, $D(t)/y(t) \rightarrow \alpha/\beta$, a (finite) constant.

EVANS PRICE ADJUSTMENT MODEL

A model of a particular market for some commodity was proposed by G. C. Evans. The demand and supply equations are the same as those for the usual simple linear model and can be solved for the equilibrium price in the usual way. In addition, there is an equation stating that the rate of change of price over time is proportional to the

excess demand ($d - s$). This factor of proportionality is positive, implying that a positive excess demand causes a rise in price and a negative excess demand causes a fall in price.

$$d(t) = \alpha_0 + \alpha_1 p(t)$$

$$s(t) = \beta_0 + \beta_1 p(t)$$

$$\frac{dp}{dt} = \gamma(d - s)$$

$$\alpha_1 < 0 \quad \beta_1 > 0 \quad \gamma > 0$$

where d is demand, s is supply, and p is price. Substituting the first two equations into the third equation,

$$\frac{dp}{dt} = \gamma[\alpha_0 - \beta_0 + (\alpha_1 - \beta_1)p]$$

$$= \gamma(\alpha_1 - \beta_1)(p - p_e)$$

where $p_e = (\alpha_0 - \beta_0)/(\beta_1 - \alpha_1)$ is the equilibrium price in the model obtained in the usual way by solving $d(t) = s(t)$ for $p(t)$, the equilibrium price. Letting $\lambda = \gamma(\alpha_1 - \beta_1)$,

$$\frac{dp}{dt} = \lambda(p - p_e)$$

$$\frac{1}{p - p_e} \frac{dp}{dt} = \lambda$$

$$\ln(p - p_e) = \lambda t + C$$

$$p - p_e = Ce^{\lambda t}$$

$$p = p_e + Ce^{\lambda t}$$

Since $p = p_0$ when $t = 0$, $C = p_0 - p_e$ and

$$p = p_e + (p_0 - p_e)e^{\lambda t}$$

where, as above, $p_e = (\alpha_0 - \beta_0)/(\beta_1 - \alpha_1)$ and $\lambda = \gamma(\alpha_1 - \beta_1)$. Since $\lambda < 0$, $p \rightarrow p_e$ as $t \rightarrow \infty$.

INCOME-CONSUMPTION-INVESTMENT MODEL

Consider a differential-equations form of the income-consumption-investment model where current consumption and investment are linear functions of current income and income changes at a rate that is proportional to excess demand, that is, to consumption plus investment minus income.

$$C_\delta(t) = \alpha Y_\delta(t)$$

$$I_\delta(t) = \gamma Y_\delta(t)$$

$$\frac{dY_\delta}{dt} = \lambda(C_\delta + I_\delta - Y_\delta)$$

$$Y_\delta(0) = Y_0 - Y_e$$

$$0 < \alpha < 1 \quad 0 < \gamma < 1 \quad 0 < \lambda < 1$$

where C_δ , I_δ , and Y_δ are deviations of consumption, investment, and income, respectively, from their equilibrium values C_e , I_e , and Y_e . Substituting the first two equations into the third equation,

$$\frac{dY_\delta}{dt} = \lambda(\alpha + \gamma - 1)Y_\delta$$

$$\frac{dY_\delta}{Y_\delta} = \lambda(\alpha + \gamma - 1) dt$$

$$\ln Y_\delta = \lambda(\alpha + \gamma - 1)t + C$$

$$Y_\delta = Ce^{\lambda(\alpha + \gamma - 1)t}$$

Since $Y_\delta = Y_0 - Y_e$ when $t = 0$, $C = Y_0 - Y_e$ and

$$Y_\delta = (Y_0 - Y_e)e^{\lambda(\alpha + \gamma - 1)t}$$

That is,

$$Y = Y_e + (Y_0 - Y_e)e^{\lambda(\alpha + \gamma - 1)t}$$

And, if $\alpha + \gamma < 1$, $e^{\lambda(\alpha + \gamma - 1)t} \rightarrow 0$ and $Y \rightarrow Y_e$ as $t \rightarrow \infty$.

PROBLEMS

1. Demand, x , and supply, y , of a product are given as a function of unit price, p , by the following equations:

$$x = ap + b \quad y = cp + d$$

Suppose that price changes in such a way that the excess of demand over supply is decreased at a rate proportional to the excess. Show that

- a. $\frac{d}{dt}(x - y) = -k(x - y)$
- b. $\frac{dp}{dt} + k(p - \bar{p}) = 0$, where $\bar{p} = \frac{b - d}{c - a}$
- c. The unit price tends to an equilibrium value \bar{p} and

$$p = \bar{p} + (p_0 - \bar{p})e^{-kt}$$

where p_0 is the initial price at $t = 0$.

2. Consider the model

$$\frac{dD}{dt} = \alpha y(t) + \beta$$

$$\frac{dy}{dt} = \gamma y(t)$$

$$y(0) = y_0$$

$$D(0) = D_0$$

$$\alpha > 0 \quad \beta > 0 \quad \gamma > 0$$

where D is national debt and y is national income.

- a. Solve the model.
- b. Determine the limit as $t \rightarrow \infty$ of the ratio of national debt to national income.

3. Consider the model

$$\frac{dR}{dt} = \frac{\alpha}{S(t)}$$

$$\frac{dS}{dt} = -\lambda S(t)$$

$$R(0) = 0$$

$$S(0) = S_0 \text{ (original cost)}$$

$$\alpha > 0 \quad \lambda > 0$$

where R is running cost of a car and S is resale value. Solve the model.

4. The increase in the number of new products tested, y , as amount, x , allocated to research and development increases equals a constant times the product of the number of new products tested and the amount allocated to research and development plus another constant times the ratio of the amount allocated to research and development divided by the number of new products tested. Find the relationship between the number of new products tested and the amount allocated to research and development if $y = y_0$ when $x = 0$.

5. Consider the model

$$\frac{dR}{dt} = \frac{\alpha}{S(t)} + \beta$$

$$\frac{dS}{dt} = -\lambda S(t)$$

$$R(0) = R_0$$

$$S(0) = S_0$$

$$\alpha > 0 \quad \beta \geq 0 \quad \lambda > 0$$

where R is the cost of repairing and operating a machine and S is salvage value of the machine. Find R and S as functions of time, t .

6. Consider the model

$$\frac{dw}{dt} = \frac{1}{\alpha y(t)} + ke^{\beta t}$$

$$\frac{dy}{dt} = \beta y(t)$$

$$w(0) = w_0$$

$$y(0) = y_0$$

$$\beta > 0 \quad \alpha > \frac{k}{\beta y_0}$$

where w is per capita consumption of wheat products and y is per capita income. Find w and y as functions of time and determine the limit as $t \rightarrow \infty$ of the ratio of per capita consumption of wheat to per capita income.

Answers to Odd-Numbered Problems

3. $S(t) = S_0 e^{-\lambda t}$

5. $S = S_0 e^{-\lambda t}$

$$R(t) = \frac{\alpha}{\lambda S_0} (e^{\lambda t} - 1)$$

$$R = R_0 + \beta t + \frac{\alpha}{\lambda S_0} (e^{\lambda t} - 1)$$

Difference Equations

6

6.1 Introduction

When variables are considered as changing discretely or discontinuously rather than continuously or instantaneously, difference equations rather than differential equations are appropriate for expressing relationships among the changes. Difference equations are frequently useful in business and economic analyses, since many economic data are recorded for uniformly spaced periods of time, for example, gross national product may be given for a year; net profit for a quarter; quantities produced, bought, or sold for a month.

In many economic analyses, time is the independent variable and changes in other variables over time are studied. Such studies of variables over discrete sets of time values are referred to as *period analyses*, and difference equations provide the basis for such analyses. Recall that when the time periods (or, more generally, the changes in the independent variable) are made smaller and smaller, difference equations approach differential equations as a limiting case.

For generality, the following discussion of the classification and solution of difference equations is in terms of an independent variable x . In many economic and business analyses, including the models discussed later, the independent variable time is denoted by t .

After the definition and classification of difference equations are discussed, the solution of a difference equation is defined. The method of solution and the behavior of the solution sequence are considered for linear first-order difference equations with constant coefficients and applications of these difference equations in economic models are illustrated. The methods of solution, the behavior of the solution sequence, and applications in economic models are then discussed for second-order difference equations.

6.2 Definition and Classification of Difference Equations

In this section, a difference equation is defined and the classification of difference equations by order and degree is discussed. Since most difference equations used in applications are linear, particular attention is given to the definition and classification of linear difference equations.

Suppose that y is a function of x , $y = f(x)$, where y is defined for integer values of x , $x = 0, 1, 2, 3, \dots$. In the context of difference equations, the functional relationship $y = f(x)$ is frequently indicated by y_x . The change in y as x changes from x to $x + 1$ is the *first difference* of y_x and is written

$$\Delta y_x = y_{x+1} - y_x \quad (\text{read "delta" } y_x)$$

Note that Δy_x is also a function of x ; Δ is an *operator* and provides the rule for computing Δy_x from the sequence y_0, y_1, y_2, \dots . Similarly, higher-order differences are obtained as differences of differences by applying the operator Δ :

$$\begin{aligned} \text{the second difference of } y_x \text{ is } \Delta^2 y_x &= \Delta(\Delta y_x) = \Delta y_{x+1} - \Delta y_x \\ &= (y_{x+2} - y_{x+1}) - (y_{x+1} - y_x) \\ &= y_{x+2} - 2y_{x+1} + y_x \end{aligned}$$

$$\begin{aligned} \text{the third difference of } y_x \text{ is } \Delta^3 y_x &= \Delta(\Delta^2 y_x) = \Delta y_{x+2} - 2\Delta y_{x+1} + \Delta y_x \\ &= (y_{x+3} - y_{x+2}) - 2(y_{x+2} - y_{x+1}) \\ &\quad + (y_{x+1} - y_x) \\ &= y_{x+3} - 3y_{x+2} + 3y_{x+1} - y_x \end{aligned}$$

$$\vdots$$

$$\text{the } k\text{th difference of } y_x \text{ is } \Delta^k y_x = \Delta(\Delta^{k-1} y_x) = \sum_{i=0}^k \frac{k!}{(k-i)! i!} (-1)^i y_{x+k-i}$$

EXAMPLE

If $y = 2x^2 - 3$,

$$\begin{aligned} \Delta^2 y &= \Delta y_{x+1} - \Delta y_x \\ &= [2(x+2)^2 - 3] - [2(x+1)^2 - 3] - \{[2(x+1)^2 - 3] - [2x^2 - 3]\} \\ &= 4 \end{aligned}$$

Alternatively,

$$\begin{aligned} \Delta^2 y_x &= y_{x+2} - 2y_{x+1} + y_x = [2(x+2)^2 - 3] - 2[2(x+1)^2 - 3] + [2x^2 - 3] \\ &= 4 \end{aligned}$$

EXAMPLE

If $y = 3x + 5$,

$$\begin{aligned}
\Delta^3 y_x &= \Delta y_{x+2} - 2\Delta y_{x+1} + \Delta y_x \\
&= [3(x+3) + 5] - [3(x+2) + 5] - 2\{[3(x+2) + 5] - [3(x+1) + 5]\} \\
&\quad + [3(x+1) + 5] - [3x + 5] \\
&= 0
\end{aligned}$$

Alternatively,

$$\begin{aligned}
\Delta^3 y_x &= y_{x+3} - 3y_{x+2} + 3y_{x+1} - y_x \\
&= [3(x+3) + 5] - 3[3(x+2) + 5] + 3[3(x+1) + 5] - [3x + 5] \\
&= 0
\end{aligned}$$

A difference equation states a relationship involving differences or, equivalently, a relationship involving the values of a dependent variable for a discrete set of (lagged) values of the independent variable. In the following discussion, the values of the independent variable are assumed to be equally spaced; for convenience, the independent variable is redefined, if necessary, so that its values are successive positive integers.

LINEAR DIFFERENCE EQUATIONS

A difference equation is said to be *linear* if the dependent variable occurs only in expressions of the first degree, that is, if the dependent variable does not occur raised to a power higher than one and does not occur in cross-product terms. Thus a difference equation is linear if it can be written in the form

$$a_0(x)y_{x+n} + a_1(x)y_{x+n-1} + \cdots + a_{n-1}(x)y_{x+1} + a_n(x)y_x = g(x)$$

where $a_0, a_1, \dots, a_{n-1}, a_n$ and g are functions of x (but not of y_x) defined for $x = 0, 1, 2, \dots$

A linear difference equation is of *order* n if, when written in this form, both a_0 and a_n are unequal to zero for all values of x under consideration; that is, a linear difference equation is of order n if it involves values of y corresponding to values of x differing by n but by not more than n . Difference equations that are not linear are, in general, very difficult to solve and are seldom used. The following discussion is confined to linear difference equations.

EXAMPLE

Each of the following difference equations is linear and of the order indicated.

$$y_{x+2} - 7y_{x+1} + 5y_x = 3x \quad \text{order 2}$$

$$3y_{x+2} + 4y_{x+1} = 2x \quad \text{order 1}$$

$$18y_{x+2} - 6y_x = 5x \quad \text{order 2}$$

$$8^x y_{x+3} - 3^x y_{x+2} + 9^x y_{x+1} + 2y_x = 3 \quad \text{order 3}$$

An n th-order difference equation can be written as an implicit function of the values of the variable y at n different values of x , that is, n lagged values of y ,

$$f[y_{x+n}, y_{x+n-1}, \dots, y_x] = 0$$

Alternatively, since knowledge of $n + 1$ adjacent values of y permits calculation of the value of y and its first n differences, an n th-order difference equation can also be written as a function of y and its first n differences,

$$F[\Delta^n y_x, \Delta^{n-1} y_x, \dots, \Delta y_x, y_x] = 0$$

EXAMPLE

$\Delta y_x = 0$ can be written

$$y_{x+1} - y_x = 0 \quad (\text{order } 1)$$

$\Delta y_x - 2y_x = 5$ can be written

$$y_{x+1} - y_x - 2y_x = 5$$

$$y_{x+1} - 3y_x - 5 = 0 \quad (\text{order } 1)$$

$\Delta^2 y_x - 3\Delta y_x - 3y_x = x$ can be written

$$\Delta(y_{x+1} - y_x) - 3(y_{x+1} - y_x) - 3y_x = x$$

$$y_{x+2} - y_{x+1} - y_{x+1} + y_x - 3y_{x+1} + 3y_x - 3y_x = x$$

$$y_{x+2} - 5y_{x+1} + y_x - x = 0 \quad (\text{order } 2)$$

$\Delta^3 y_x + \Delta^2 y_x + \Delta y_x + y_x = 0$ can be written

$$\Delta^2(y_{x+1} - y_x) + \Delta(y_{x+1} - y_x) + y_{x+1} - y_x + y_x = 0$$

$$\Delta(y_{x+2} - y_{x+1} - y_{x+1} + y_x) + y_{x+2} - y_{x+1} - y_{x+1} + y_x + y_{x+1} = 0$$

$$y_{x+3} - y_{x+2} - 2y_{x+2} + 2y_{x+1} + y_{x+1} - y_x + y_{x+2} - y_{x+1} + y_x = 0$$

$$y_{x+3} - 2y_{x+2} + 2y_{x+1} = 0 \quad (\text{order } 2)$$

6.3 Solutions of Difference Equations

A *solution* of a difference equation is a functional relation, not involving differences, which is defined for all nonnegative integers and which satisfies the difference equation.

The *general solution* of an n th-order difference equation is a solution containing n arbitrary constants. It can be shown that the general solution of a difference equation is unique.

A *particular solution* of a difference equation is a solution that can be obtained from the general solution by giving particular values to the arbitrary constants of the general solution. As in the case of differential equations, the arbitrary constants of a difference equation are specified by *boundary conditions* or *initial conditions*. The general solution of an n th-order difference equation includes n arbitrary constants, and a particular solution thus requires the specification of n boundary conditions.

In each of the following examples a proposed solution is shown to satisfy a given difference equation and a particular solution is determined for the specified boundary condition(s). Methods for obtaining solutions of linear first- and second-order difference equations with constant coefficients are then discussed. Similar methods of solution are appropriate for higher-order linear difference equations.

EXAMPLE

Show that $y_x = x + c$ is a solution of $y_{x+1} - y_x = 1$ and find a particular solution if $y_0 = 1$.

If $y_x = x + c$,

$$y_{x+1} - y_x = 1$$

$$(x + 1 + c) - (x + c) = 1$$

so

$$y_x = x + c$$

is a solution.

If $y_0 = 1$,

$$1 = 0 + c$$

$$c = 1$$

and

$$y_x = x + 1$$

is the particular solution when $y_0 = 1$.

EXAMPLE

Show that $y_x = [x(x-1)/2] + c$ is a solution of $y_{x+1} - y_x = x$ and find a particular solution if $y_0 = 2$.

If $y_x = x(x-1)/2$,

$$y_{x+1} - y_x = x$$

$$\frac{(x+1)x}{2} + c - \left(\frac{x(x-1)}{2} + c \right) = x$$

$$\frac{x(x+1-x+1)}{2} = x$$

$$\frac{2x}{2} = x$$

so

$$y_x = \frac{x(x-1)}{2} + c$$

is a solution.

If $y_0 = 2$,

$$2 = 0 + c$$

$$c = 2$$

and

$$y_x = \frac{x(x-1)}{2} + 2$$

is the particular solution when $y_0 = 2$.

EXAMPLE

Show that $y_x = c_1 + c_2(-1)^x$ is a solution of $y_{x+2} - y_x = 0$ and find a particular solution if $y_0 = 2, y_1 = 5$.

If $y_x = c_1 + c_2(-1)^x$,

$$y_{x+2} - y_x = 0$$

$$c_1 + c_2(-1)^{x+2} - [c_1 + c_2(-1)^x] = 0$$

$$c_2[(-1)^{x+2} - (-1)^x] = 0$$

so

$$y_x = c_1 + c_2(-1)^x$$

is a solution.

If $y_0 = 2, y_1 = 5$,

$$2 = c_1 + c_2$$

$$5 = c_1 - c_2$$

$$2c_1 = 7$$

$$c_1 = \frac{7}{2}$$

$$c_2 = -\frac{3}{2}$$

and

$$y_x = \frac{7}{2} - \frac{3}{2}(-1)^x$$

is the particular solution for $y_0 = 2, y_1 = 5$.

Analogies Between Difference Equations and Differential Equations

There are many analogies in the classification and solution of difference and differential equations; these arise because the derivative of a function is defined as a limit of a difference quotient

$$\frac{dy}{dx} = \lim_{\Delta x \rightarrow 0} \frac{\Delta y_x}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{y_{x+\Delta x} - y_x}{\Delta x}$$

and, as $\Delta x \rightarrow 0$, the discrete case (difference equations) approaches as a limit the continuous case (differential equations).

PROBLEMS

1. Determine the order of each of the following linear difference equations.

a. $3y_{x+2} - 3y_{x+1} = 3x$

c. $7y_{x+1} - 5y_x = 3$

b. $8y_{x+3} - y_x = 4$

d. $6y_{x+2} - 7y_x = 5x$

2. If $y = x^2 + 2x$, find $\Delta^2 y_x$.

3. If $y = e^x$, find $\Delta^2 y_x$.

4. If $y = x^3 + 3$, find $\Delta^2 y_x$ and $\Delta^3 y_x$.

5. Show that $y_x = \frac{c}{1+cx}$ is a solution of $y_{x+1} = \frac{y_x}{1+y_x}$ and find a particular solution if $y_0 = -4$.

6. Show that $y_x = c_1 + c_2 2^x$ is a solution of $y_{x+2} - 3y_{x+1} + 2y_x = 0$ and find a particular solution if $y_0 = 1, y_1 = 2$.

7. Show that $y_x = c_1 + c_2 2^x - x$ is a solution of $y_{x+2} - 3y_{x+1} + 2y_x = 1$ and find a particular solution if $y_0 = 0, y_1 = 3$.
8. Show that $y_x = c_1 + c_2 2^x + c_3 3^x$ is a solution of $y_{x+3} - 6y_{x+2} + 11y_{x+1} - 6y_x = 0$ and find a particular solution if $y_0 = 1, y_1 = 1, y_2 = -1$.
9. Show that $y_x = c_1 + c_2 x + c_3 x^2 + c_4 x^3$ is a solution of $y_{x+4} - 4y_{x+3} + 6y_{x+2} - 4y_{x+1} + y_x = 0$ and find a particular solution if $y_0 = 1, y_1 = 5, y_2 = 9, y_3 = 7$.
10. Write each of the following difference equations in terms of lagged values of y .
 - a. $\Delta y_x = 10$
 - b. $\Delta^2 y_x - 3\Delta y_x - 5 = 0$
 - c. $\Delta^2 y_x - 4y_x = 2$
 - d. $\Delta^3 y_x + 5\Delta y_x = y_x$

Answers to Odd-Numbered Problems

1. a. order 1
b. order 3
c. order 1
d. order 2
3. $e^x(e-1)^2$
5. $y_x = \frac{4}{4x-1}$
7. $y_x = 4(2^x - 1) - x$
9. $y_x = 1 + 2x + 3x^2 - x^3$

6.4 Linear First-Order Difference Equations with Constant Coefficients

In this section, the method of solution for linear first-order difference equations with constant coefficients is discussed. The behavior of the solution sequence is considered and equilibrium and stability are defined.

A linear first-order difference equation can be written in the form

$$a_0(x)y_{x+1} + a_1(x)y_x = g(x) \quad x = 0, 1, 2, \dots$$

where $a_0(x) \neq 0$ and $a_1(x) \neq 0$ or, alternatively, as

$$y_{x+1} = -\frac{a_1(x)}{a_0(x)}y_x + \frac{g(x)}{a_0(x)}$$

If $a_0(x)$, $a_1(x)$, and $g(x)$ are constants, that is, are not in fact functions of x , then

$$y_{x+1} = Ay_x + B$$

where A and B are constants and $A \neq 0$. ($B = 0$ if and only if $g = 0$ in the original equation.) The difference equation $y_{x+1} = Ay_x + B$ is the general first-order difference equation with constant coefficients.

The solution of this equation can be obtained by induction as follows:

$$\begin{aligned} y_1 &= Ay_0 + B \\ y_2 &= Ay_1 = A(Ay_0 + B) + B \\ &= A^2y_0 + AB + B \\ y_3 &= Ay_2 = A(A^2y_0 + AB + B) + B \\ &= A^3y_0 + A^2B + AB + B \end{aligned}$$

$$\begin{aligned}
 y_4 &= Ay_3 = A(A^3y_0 + A^2B + AB + B) + B \\
 &= A^4y_0 + A^3B + A^2B + AB + B \\
 &\vdots \\
 y_x &= A^xy_0 + B(1 + A + A^2 + A^3 + \cdots + A^{x-1})
 \end{aligned}$$

and thus, noting that $1 + A + A^2 + A^3 + \cdots + A^{x-1}$ is a geometric series with sum $(1 - A^x)/(1 - A)$, the solution of $y_{x+1} = Ay_x + B$ is

$$y_x = A^xy_0 + B \frac{1 - A^x}{1 - A} \quad \text{for } A \neq 1 \quad x = 0, 1, 2, \dots$$

$$y_x = y_0 + Bx \quad \text{for } A = 1 \quad x = 0, 1, 2, \dots$$

This solution, obtained by induction, does in fact satisfy the equation $y_{x+1} = Ay_x + B$, since

For $A \neq 1$,

$$\begin{aligned}
 y_{x+1} &= Ay_x + B \\
 &= A \left(A^xy_0 + B \frac{1 - A^x}{1 - A} \right) + B \\
 &= A^{x+1}y_0 + B \left(\frac{A - A^{x+1} + 1 - A}{1 - A} \right) \\
 &= A^{x+1}y_0 + B \frac{1 - A^{x+1}}{1 - A}
 \end{aligned}$$

For $A = 1$,

$$\begin{aligned}
 y_{x+1} &= y_x + B \\
 &= (y_0 + Bx) + B \\
 &= y_0 + B(x + 1)
 \end{aligned}$$

There are three special cases of the difference equation $y_{x+1} = Ay_x + B$ which occur frequently in descriptions of business and economic data.

1. First-order difference a constant:

$$y_{x+1} - y_x = B \quad (\text{special case: } A = 1)$$

$$\text{Solution: } y_x = y_0 + Bx$$

2. First-order difference proportional to the variable:

$$y_{x+1} - y_x = \alpha y_{x+1} \quad \left(\text{special case: } A = \frac{1}{1 - \alpha}, B = 0 \right)$$

$$\text{Solution: } y_x = \left(\frac{1}{1 - \alpha} \right)^x y_0$$

3. First-order difference a linear function of the variable:

$$y_{x+1} - y_x = \alpha y_{x+1} + \beta \quad \left(\text{special case: } A = \frac{1}{1 - \alpha}, B = \frac{1}{1 - \alpha} \right)$$

$$\text{Solution: } y_x = \left(\frac{1}{1 - \alpha} \right)^x y_0 + \frac{\beta}{\alpha} \left[\left(\frac{1}{1 - \alpha} \right)^x - 1 \right]$$

Examples of solutions of linear first-order difference equations with constant coefficients and of economic models in which they occur are given in the following sections.

EXAMPLE

Solve the equation $2y_{x+1} = 4y_x + 3$ and find the particular solution if $y_0 = \frac{1}{2}$.

$$2y_{x+1} = 4y_x + 3$$

$$y_{x+1} = 2y_x + \frac{3}{2}$$

$$y_x = 2^x y_0 + \frac{3}{2} \left(\frac{1 - 2^x}{-1} \right)$$

$$y_x = 2^x y_0 - \frac{3}{2}(1 - 2^x)$$

$$y_x = (y_0 + \frac{3}{2})2^x - \frac{3}{2}$$

If $y_0 = \frac{1}{2}$,

$$y_x = 2^{x+1} - \frac{3}{2}$$

EXAMPLE

Solve the equation $3y_{x+1} = 3y_x - 7$ and find the particular solution if $y_0 = 3$.

$$3y_{x+1} = 3y_x - 7$$

$$y_{x+1} = y_x - \frac{7}{3}$$

$$y_x = y_0 - \frac{7}{3}x$$

If $y_0 = 3$,

$$y_x = 3 - \frac{7}{3}x$$

EXAMPLE

Solve the equation $6y_{x+1} + 2y_x = 0$ and find the particular solution if $y_0 = 2$.

$$6y_{x+1} + 2y_x = 0$$

$$y_{x+1} = -\frac{1}{3}y_x$$

$$y_x = (-\frac{1}{3})^x y_0$$

If $y_0 = 2$,

$$y_x = 2(-\frac{1}{3})^x$$

EXAMPLE

Solve the equation $3y_{x+1} - 9y_x + 8 = 0$ and find the particular solution if $y_0 = \frac{1}{3}$.

$$3y_{x+1} - 9y_x + 8 = 0$$

$$y_{x+1} = 3y_x - \frac{8}{3}$$

$$y_x = 3^x y_0 - \frac{8}{3} \left(\frac{1 - 3^x}{-2} \right)$$

$$y_x = 3^x y_0 + \frac{4}{3}(1 - 3^x)$$

$$y_x = (y_0 - \frac{4}{3})3^x + \frac{4}{3}$$

$$\text{If } y_0 = \frac{1}{3},$$

$$y_x = \frac{4}{3} - 3^x$$

BEHAVIOR OF THE SOLUTION SEQUENCE

A sequence is a succession of terms formed according to some rule or, equivalently, a sequence is a function defined for positive integer values of the independent variable. The solution of a difference equation is a sequence. When the independent variable is time, this sequence is sometimes referred to as the time path of the dependent variable. For a linear first-order difference equation, the specification of y_0 determines or generates the solution sequence $y_0, y_1, y_2, y_3, \dots$; each term is found from the difference equation

$$y_{x+1} = Ay_x + B \quad x = 0, 1, 2, \dots$$

or, equivalently, from its solution,

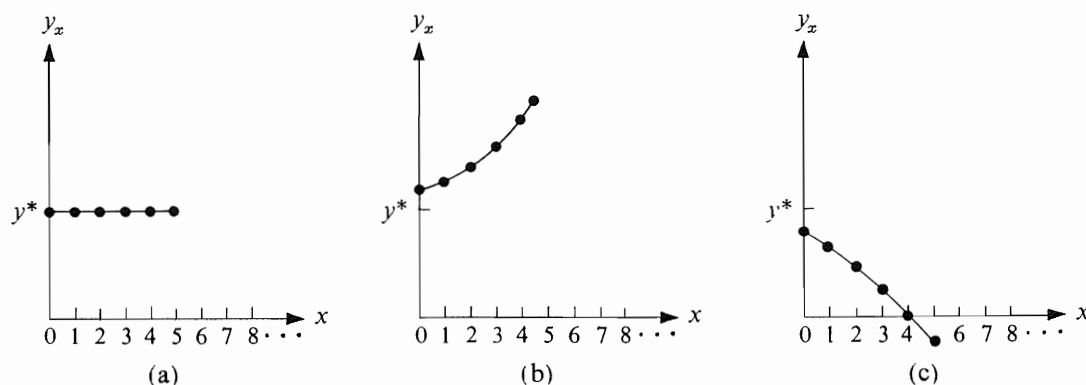
$$y_x = A^x y_0 + B \frac{1 - A^x}{1 - A} \quad \text{for } A \neq 1 \quad x = 0, 1, 2, \dots$$

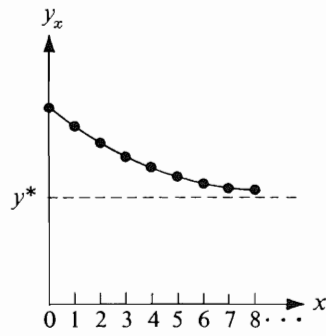
$$y_x = y_0 + Bx \quad \text{for } A = 1 \quad x = 0, 1, 2, \dots$$

When A and B are given, specification of y_0 thus determines a solution sequence of real numbers.

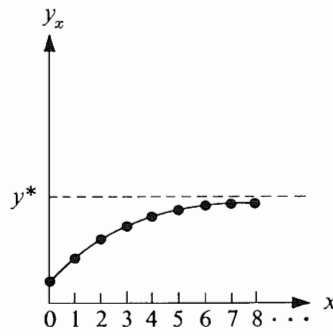
The behavior of the solution sequence of a difference equation is of interest for many applications; this behavior depends on the values of y_0 , A , and B , as shown in Table 6.1. A sketch typical of each type of behavior is given in Figure 6.1, labeled to match the corresponding case in the table.

Figure 6.1

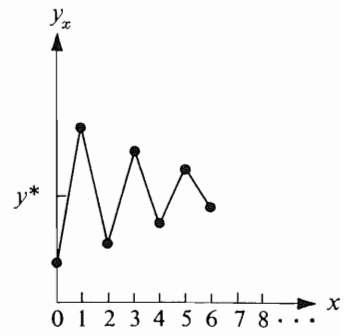




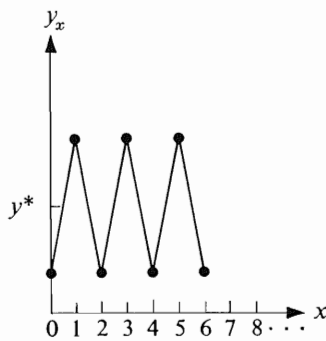
(d)



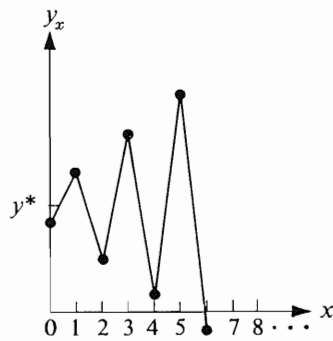
(e)



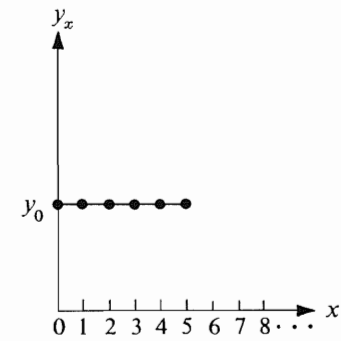
(f)



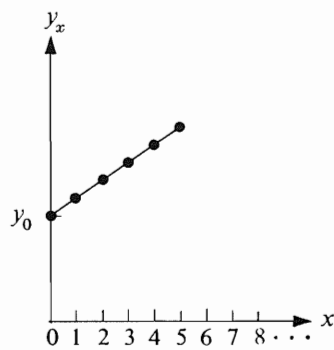
(g)



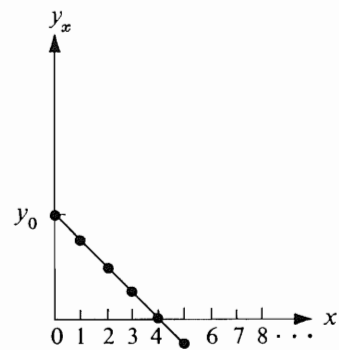
(h)



(i)



(j)



(k)

Table 6.1 Behavior of the Solution Sequence for $y_{x+1} = Ay_x + B$

Case	A	B	y_0	$y_x, x = 1, 2, \dots$	Behavior of the solution sequence
(a)	$A \neq 1$		$y_0 = y^*$	$y_x = y^*$	Constant: $y_x = y^*$
(b)	$A > 1$		$y_0 > y^*$	$y_x > y^*$	Diverges to $+\infty$ (monotone increasing)
(c)	$A > 1$		$y_0 < y^*$	$y_x < y^*$	Diverges to $-\infty$ (monotone decreasing)
(d)	$0 < A < 1$		$y_0 > y^*$	$y_x > y^*$	Converges to y^* (monotone decreasing)
(e)	$0 < A < 1$		$y_0 < y^*$	$y_x < y^*$	Converges to y^* (monotone increasing)
(f)	$-1 < A < 0$		$y_0 \neq y^*$		Converges to y^* (damped oscillatory)
(g)	$A = -1$		$y_0 \neq y^*$		Diverges (oscillates finitely)
(h)	$A < -1$		$y_0 \neq y^*$		Diverges (oscillates infinitely)
(i)	$A = 1$	$B = 0$		$y_x = y_0$	Constant: $y_x = y_0$
(j)	$A = 1$	$B > 0$		$y_x > y_0$	Diverges to $+\infty$ (monotone increasing)
(k)	$A = 1$	$B < 0$		$y_x < y_0$	Diverges to $-\infty$ (monotone decreasing)

The results given in the table can be summarized in the following theorem.

THEOREM: The linear first-order difference equation

$$y_{x+1} = Ay_x + B \quad x = 0, 1, 2, \dots$$

has the solution

$$y_x = A^x(y_0 - y^*) + y^* \quad \text{if } A \neq 1 \quad x = 0, 1, 2, \dots$$

$$y_x = y_0 + Bx \quad \text{if } A = 1 \quad x = 0, 1, 2, \dots$$

where $y^* = B/(1 - A)$. If $-1 < A < 1$, the solution sequence converges to y^* ; otherwise it diverges, unless $y_x = y_0$.

EXAMPLE

Solve each of the following difference equations, determine the behavior of the solution sequence, and calculate the first few values of the solution sequence.

(a) $6y_{x+1} + 2y_x - 3 = 0, y_0 = 1.$

$$y_{x+1} = -\frac{1}{3}y_x + \frac{1}{2}$$

$$y_x = \left(-\frac{1}{3}\right)^x y_0 + \frac{1}{2} \left(\frac{1 - \left(-\frac{1}{3}\right)^x}{1 - \left(-\frac{1}{3}\right)} \right)$$

$$= \left(-\frac{1}{3}\right)^x y_0 + \frac{3}{8} (1 - \left(-\frac{1}{3}\right)^x)$$

$$= \left(-\frac{1}{3}\right)^x (y_0 - \frac{3}{8}) + \frac{3}{8}$$

If $y_0 = 1,$

$$y_x = \frac{5}{8} \left(-\frac{1}{3}\right)^x + \frac{3}{8}$$

$A = -\frac{1}{3}$, $B = \frac{1}{2}$, $y^* = \frac{3}{8}$; case (f), so y_x converges to y^* , damped oscillatory. $y_0 = 1$, $y_1 = \frac{1}{6}$, $y_2 = \frac{4}{9}$, $y_3 = \frac{19}{54}$, $y_4 = \frac{31}{81}$, and so forth.

(b) $5y_{x+1} - 4y_x - 15 = 0$, $y_0 = 10$.

$$y_{x+1} = \frac{4}{5}y_x + 3$$

$$\begin{aligned} y_x &= \left(\frac{4}{5}\right)^x y_0 + 3 \left(\frac{1 - \left(\frac{4}{5}\right)^x}{1 - \frac{4}{5}} \right) \\ &= \left(\frac{4}{5}\right)^x y_0 + 15 \left(1 - \left(\frac{4}{5}\right)^x \right) \\ &= \left(\frac{4}{5}\right)^x (y_0 - 15) + 15 \end{aligned}$$

If $y_0 = 10$,

$$y_x = 15 - 5\left(\frac{4}{5}\right)^x$$

$A = \frac{4}{5}$, $B = 3$, $y^* = 15$; case (d), so y_x converges to y^* , monotone increasing. $y_0 = 10$, $y_1 = 11$, $y_2 = 11\frac{4}{5}$, $y_3 = 12\frac{11}{25}$, $y_4 = 12\frac{119}{625}$, and so forth.

(c) $5y_{x+1} - 25y_x - 1 = 0$, $y_0 = 0$.

$$y_{x+1} = 5y_x + \frac{1}{5}$$

$$\begin{aligned} y_x &= 5^x y_0 + \frac{1}{5} \left(\frac{1 - 5^x}{1 - 5} \right) \\ &= 5^x y_0 - \frac{1}{20} (1 - 5^x) \\ &= 5^x \left(y_0 + \frac{1}{20} \right) - \frac{1}{20} \end{aligned}$$

If $y_0 = 0$,

$$y_x = \frac{1}{20} (5^x - 1)$$

$A = 5$, $B = \frac{1}{5}$, $y^* = -\frac{1}{20}$; case (b), so y_x diverges to $+\infty$, monotone increasing. $y_0 = 0$, $y_1 = \frac{1}{5}$, $y_2 = \frac{6}{5}$, $y_3 = \frac{31}{5}$, $y_4 = \frac{156}{5}$, and so forth.

EXAMPLE

If money earns simple interest, the amount on deposit at any interest date is equal to the amount on deposit one period (usually a year) previously plus the interest earned in that period on the initial principal invested. If S_0 is the initial deposit, S_t is the amount on deposit after t periods, and r is the interest rate per period, then

$$S_{t+1} = S_t + rS_0 \quad t = 0, 1, 2, \dots$$

which has the solution

$$S_t = S_0(1 + tr) \quad t = 0, 1, 2, \dots$$

When the initial amount accumulates at compound interest, the interest earned during any period is computed on the total sum on deposit at the beginning of that period, rather than on the initial deposit, and

$$S_{t+1} = S_t + iS_t = (1 + i)S_t \quad t = 0, 1, 2, \dots$$

where i is the interest rate per conversion period; thus

$$S_t = (1 + i)^t S_0 \quad t = 0, 1, 2, \dots$$

The amount on deposit diverges to $+\infty$ linearly for simple interest [case (j)] and more rapidly for compound interest [case (b)].

EQUILIBRIUM AND STABILITY

If a difference equation has a constant function $y_x = y^*$ as its solution, then the value y^* of this function is called an *equilibrium* or *stationary* value of y_x . An equilibrium value y^* , or the corresponding difference equation, is said to be *stable* if every solution of the difference equation converges to y^* , independent of the initial conditions. This is the type of stability referred to in economic literature as "perfect stability of the first kind."

A displacement from the equilibrium value is equivalent to a new solution with different initial conditions, and thus a stable equilibrium can be defined as one for which *any* displacement from equilibrium is followed by a sequence of values of y which again converge to equilibrium.

THEOREM: For the difference equation

$$y_{x+1} = Ay_x + B$$

an equilibrium value of y is given by

$$y^* = \frac{B}{1-A} \quad \text{if } A \neq 1$$

and y^* is stable if and only if $-1 < A < 1$, unless y_x is constant.

Note that this theorem is closely related to the preceding theorem, which states that the solution sequence converges to y^* if $-1 < A < 1$ and diverges otherwise, unless $y_x = y_0$. Stable equilibrium occurs only for cases (a), (d), (e), (f), and (i) of Table 6.1.

PROBLEMS

Solve each of the following difference equations.

3. $y_{x+1} + 3y_x = 0$
4. $y_{x+1} + y_x - 2 = 0$
1. $3y_{x+1} = 2y_x + 3$
2. $2y_{x+1} + y_x - 3 = 0$

Solve each of the following difference equations, determine the behavior of the solution sequence, and calculate the first few values of the solution sequence.

5. $y_{x+1} - y_x - 10 = 0, y_0 = 2$
6. $y_{x+1} = 7y_x + 6, y_0 = 1$
7. $8y_{x+1} + 4y_x - 3 = 0, y_0 = \frac{1}{2}$
8. $16y_{x+1} - 6y_x = 1, y_0 = \frac{1}{10}$
9. $3y_{x+1} - 2y_x - 3 = 0, y_0 = 5$
10. $3y_{x+1} - y_x = \frac{6}{5}, y_0 = \frac{2}{5}$
11. $y_{x+1} = 3y_x - 1, y_0 = \frac{1}{2}$
12. $y_{x+1} + 3y_x + 1 = 0, y_0 = 1$
13. $2y_{x+1} - y_x = 2, y_0 = 4$
14. $y_{x+1} = y_x - 1, y_0 = 5$
15. $7y_{x+1} + 2y_x - 7 = 0, y_0 = 1$
16. $y_{x+1} + y_x + 2 = 0, y_0 = 3$
17. $15y_{x+1} - 10y_x - 3 = 0, y_0 = 1$
18. $5y_{x+1} - y_x - 60 = 0, y_0 = 15$
19. $y_{x+1} + 4y_x + 12 = 0, y_0 = 6$
20. $8y_{x+1} + y_x - 4 = 0, y_0 = \frac{1}{3}$
21. $4y_{x+1} - y_x - 3 = 0, y_0 = \frac{1}{2}$
22. $4y_{x+1} + 3y_x - 4 = 0, y_0 = 1$
23. $3y_{x+1} - 2y_x - 6 = 0, y_0 = 4$
24. $9y_{x+1} + 5y_x - 18 = 0, y_0 = 1$

Answers to Odd-Numbered Problems

1. $y_x = (\frac{2}{3})^x(y_0 - 3) + 3$
3. $y_x = (-3)^x y_0$
5. $y_x = 2 + 10x$; monotone increasing, diverges to $+\infty$; $y_0 = 2, y_1 = 12, y_2 = 22, y_3 = 32, y_4 = 42, y_5 = 52$
7. $y_x = \frac{1}{4}(-\frac{1}{2})^x + \frac{1}{4}$; damped oscillatory, converges to $y^* = \frac{1}{4}$; $y_0 = \frac{1}{2}, y_1 = \frac{1}{8}, y_2 = \frac{5}{16}, y_3 = \frac{3}{32}, y_4 = \frac{17}{64}, y_5 = \frac{31}{128}$
9. $y_x = 2(\frac{2}{3})^x + 3$; monotone decreasing, converges to $y^* = 3$; $y_0 = 5, y_1 = 4\frac{1}{3}, y_2 = 3\frac{8}{9}, y_3 = 3\frac{16}{27}, y_4 = 3\frac{32}{81}, y_5 = 3\frac{64}{243}$

11. $y_x = \frac{1}{2}$; constant
13. $y_x = 2(\frac{1}{2})^x + 2$; monotone decreasing, converges to $y^* = 2$; $y_0 = 4$, $y_1 = 3$, $y_2 = \frac{5}{2}$, $y_3 = \frac{9}{4}$,
 $y_4 = \frac{17}{8}$, $y_5 = \frac{33}{16}$
15. $y_x = \frac{2}{3}(-\frac{2}{3})^x + \frac{7}{3}$; damped oscillatory, converges to $y^* = \frac{7}{3}$; $y_0 = 1$, $y_1 = \frac{5}{3}$, $y_2 = \frac{39}{49}$,
 $y_3 = \frac{265}{343}$, $y_4 = \frac{1871}{2401}$, $y_5 = \frac{13065}{16807}$
17. $y_x = \frac{2}{3}(\frac{2}{3})^x + \frac{3}{5}$; monotone decreasing, converges to $y^* = \frac{3}{5}$; $y_0 = 1$, $y_1 = \frac{13}{15}$, $y_2 = \frac{7}{9}$,
 $y_3 = \frac{97}{135}$, $y_4 = \frac{55}{81}$, $y_5 = \frac{29}{45}$
19. $y_x = \frac{4}{5}(-4)^x - \frac{1}{5}$; divergent, oscillates infinitely; $y_0 = 6$, $y_1 = -36$, $y_2 = 132$,
 $y_3 = -540$, $y_4 = 2148$, $y_5 = -8604$
21. $y_x = -(\frac{1}{2})(\frac{1}{4})^x + 1$; monotone increasing, converges to $y^* = 1$; $y_0 = \frac{1}{2}$, $y_1 = \frac{7}{8}$, $y_2 = \frac{31}{32}$,
 $y_3 = \frac{127}{128}$, $y_4 = \frac{511}{512}$, $y_5 = \frac{2047}{2048}$
23. $y_x = -2(\frac{2}{3})^x + 6$; monotone increasing, converges to $y^* = 6$; $y_0 = 4$, $y_1 = \frac{14}{3}$, $y_2 = \frac{46}{9}$,
 $y_3 = \frac{146}{27}$, $y_4 = \frac{454}{81}$, $y_5 = \frac{1394}{243}$

6.5 Applications of Difference Equations in Economic Models

The following very simple economic models are examples of the use of first-difference equations to express the dependence of the value of a variable on its value in the preceding period. The Harrod model, the general cobweb model, a consumption model, and an income-consumption-investment model are discussed.

THE HARROD MODEL

The following model for national income analysis was proposed by R. F. Harrod.

$$S_t = \alpha Y_t$$

$$I_t = \beta(Y_t - Y_{t-1})$$

$$S_t = I_t$$

$$Y_0 = Y_0 \text{ (known value at } t = 0\text{)}$$

$$\alpha > 0 \quad \beta > 0$$

where S is saving, I is investment, Y is income, and each of these variables is a function of time t .

From the first three equations of the model, the difference equation

$$\alpha Y_t = \beta(Y_t - Y_{t-1})$$

is obtained for solution and

$$Y_t = \left(\frac{\beta}{\beta - \alpha} \right) Y_{t-1}$$

$$Y_t = \left(\frac{\beta}{\beta - \alpha} \right)^t Y_0$$

$$I_t = S_t = \alpha \left(\frac{\beta}{\beta - \alpha} \right)^t Y_0$$

Assuming $Y_0 > 0$, the behavior of the solution sequence depends on the value of the constant $\beta/(\beta - \alpha)$. Since Y represents income and is assumed to be nonnegative,

$$\frac{\beta}{\beta - \alpha} \geq 0$$

and from the model, $\alpha > 0$, $\beta > 0$, and thus

$$\frac{\beta}{\beta - \alpha} > 1$$

Since $Y^* = 0$, the sequence $\{Y_t\}$ is monotone increasing, diverging to $+\infty$ [case (b)]. Since $\alpha > 0$, the sequences $\{I_t\}$ and $\{S_t\}$ are also monotone increasing, diverging to $+\infty$. None of the variables of this model have equilibrium values.

NOTE: The Harrod model above is one of the classical models used to study the growth of national income in an expanding economy; models for which the solution sequences do not diverge have also been considered.

THE GENERAL COBWEB MODEL

The adjustment of supply and demand can be studied using the following model.

$$\text{Supply: } q_t = \alpha + \beta p_{t-1}$$

$$\text{Demand: } p_t = \gamma + \delta q_t$$

$$q_0 = q_0 \quad (\text{known value at } t = 0)$$

where p is price, q is quantity, and both are functions of time t . Substituting the second equation into the first equation,

$$q_t = \alpha + \beta\gamma + \beta\delta q_{t-1}$$

and, substituting the first equation into the second equation,

$$p_t = \gamma + \alpha\delta + \beta\delta p_{t-1}$$

Difference equations can be solved in the usual way; however, it is the behavior of the solution sequence that is of interest. Since $\beta > 0$, $\delta < 0$, their product $\beta\delta < 0$, and the solution sequences are thus always oscillatory. The equilibrium point is

$$(p^*, q^*) = \left(\frac{\gamma + \alpha\delta}{1 - \beta\delta}, \frac{\alpha + \beta\gamma}{1 - \beta\delta} \right)$$

If $-1 < \beta\delta < 0$, the sequences $\{p_t\}$ and $\{q_t\}$ are damped and converge to (p^*, q^*) ; if $\beta\delta = -1$, the sequences oscillate finitely; if $\beta\delta < -1$, the sequences oscillate infinitely. Thus the equilibrium is stable only if $-1 < \beta\delta < 0$.

NOTE: If the supply curve is written so that its slope is with respect to the quantity variable, the slope is $1/\beta$. Thus the equilibrium is stable if and only if the slope of the demand curve is less (in numerical value) than the slope of the supply curve.

CONSUMPTION MODEL

The following simple model for consumption has been proposed.

$$C_t + S_t = Y_t$$

$$Y_t = \alpha S_{t-1}$$

$$C_t = \gamma Y_t$$

$$Y_0 = Y_0 \quad (\text{known value at } t = 0)$$

$$\alpha > 0 \quad 0 < \gamma < 1$$

where C is consumption, S is saving, Y is income, and each of these variables is a function of time, t , and γ is the marginal propensity to consume.

Substituting the second and third equations in the first equation of the model, the difference equation

$$Y_{t+1} = \alpha(1 - \gamma)Y_t$$

is obtained for solution and

$$Y_t = (\alpha - \alpha\gamma)^t Y_0$$

$$C_t = \gamma Y_t$$

so

$$C_t = \gamma(\alpha - \alpha\gamma)^t Y_0$$

and, since $C_0 = \gamma Y_0$,

$$C_t = (\alpha - \alpha\gamma)^t C_0$$

$$S_t = Y_t - C_t$$

so

$$S_t = (1 - \gamma)(\alpha - \alpha\gamma)^t Y_0$$

and, since $S_0 = Y_0 - C_0 = (1 - \gamma)Y_0$,

$$S_t = (\alpha - \alpha\gamma)^t S_0$$

Since $0 < \gamma < 1$, $\alpha - \alpha\gamma > 0$; the sequences $\{Y_t\}$, $\{C_t\}$, and $\{S_t\}$ are monotone increasing and diverge to $+\infty$ if $\alpha(1 - \gamma) > 1$, monotone decreasing and converge to $y^* = 0$ if $\alpha(1 - \gamma) < 1$, and constant at $Y^* = 0$ if $\alpha(1 - \gamma) = 1$.

INCOME-CONSUMPTION-INVESTMENT MODEL

When changes over time are considered to occur periodically rather than continuously, the income-consumption-investment model stated in terms of differential equations in Chapter 5 can be stated in terms of difference equations, as follows.

$$C_t = \alpha Y_t + \beta$$

$$I_t = \gamma Y_t + g$$

$$\Delta Y_{t-1} = \theta[C_{t-1} + I_{t-1} - Y_{t-1}]$$

$$Y_0 = Y_0$$

$$0 < \alpha < 1 \quad 0 < \gamma < 1 \quad 0 < \theta < 1$$

Rewriting the third equation and substituting the first and second equations into it,

$$Y_t - Y_{t-1} = \theta[C_{t-1} + I_{t-1} - Y_{t-1}]$$

$$Y_t = \theta[C_{t-1} + I_{t-1}] + (1 - \theta)Y_{t-1}$$

$$= \theta[\alpha Y_{t-1} + \beta + \gamma Y_{t-1} + g] + (1 - \theta)Y_{t-1}$$

$$= [\theta(\alpha + \gamma) + (1 - \theta)]Y_{t-1} + \theta(\beta + g)$$

$$= [\theta(\alpha + \gamma) + (1 - \theta)]^t Y_0 + \theta(\beta + g) \frac{1 - [\theta(\alpha + \gamma) + (1 - \theta)]^t}{1 - [\theta(\alpha + \gamma) + (1 - \theta)]}$$

The solution is stable if

$$-1 < \theta(\alpha + \gamma) + (1 - \theta) < 1$$

$$-\frac{1}{\theta} < \alpha + \gamma + \frac{1}{\theta} - 1 < \frac{1}{\theta}$$

$$1 - \frac{2}{\theta} < \alpha + \gamma < 1$$

Since $0 < \theta < 1$ and $\alpha + \gamma > 0$, the left-hand part of the inequality $1 - (2/\theta) < \alpha + \gamma$ is satisfied. Thus the stability condition is $\alpha + \gamma < 1$, as for the differential equations form of the model.

PROBLEMS

1. In the Cobweb model, if the supply curve has negative slope, state the conditions for stable equilibrium.
2. Derive the solution of the Kahn model:

$$C_t = \alpha Y_{t-1} + \beta$$

$$Y_t = C_t + I_t$$

where C is consumption, Y is income, and I is investment.

3. Solve the following model, determine the behavior of the solution sequences, and state any additional "logical" restrictions on the parameters.

$$S_t = \alpha Y_t + \beta$$

$$I_t = \gamma(Y_t - Y_{t-1})$$

$$S_t = \delta I_t$$

$$Y_0 = Y_0$$

$$\alpha > 0 \quad \beta > 0 \quad \gamma > 0 \quad \delta > 0$$

where S is savings, Y is income, and I is investment.

4. Solve the following simplified Metzler inventory model:

$$y_t = u_t + v_0$$

$$u_t = \beta y_{t-1}$$

$$y_0 = y_0$$

$$0 < \beta < 1$$

where y is income produced, u is the number of units produced for sale, with units of measurement appropriately chosen, v_0 is the constant noninduced investment, and β is the marginal propensity to consume of one year's consumption with respect to the previous year's income.

5. Consider the model

$$p_t = 2 - q_t$$

$$q_{t+1} = p_{t+1}^* + 1$$

$$p_{t+1}^* = \alpha \rho_{t+1} + (1 - \alpha)p_t$$

$$p_0 = p_0$$

$$0 \leq \alpha \leq 1$$

where q is the quantity supplied, p is the actual price, p^* is the supplier's estimated price for the next period, and ρ is a public forecast for the price. Note that $\alpha = 1$ implies that the

suppliers have perfect confidence in the public forecast; $\alpha = 0$ implies they have no confidence in the public forecast. Defining Δ_{t+1} by the equation

$$\Delta_{t+1} = p_{t+1}^* - p_{t+1}$$

show that $\Delta_t = 1 - 2p_t$.

6. Solve the following model of the growth of national income in an expanding economy:

$$Y_t = C_t + I_t$$

$$C_t = \alpha + \beta Y_t$$

$$Y_{t+1} - Y_t = \gamma I_t$$

$$Y_0 = Y_0 \quad C_0 = C_0 \quad I_0 = I_0$$

$$\alpha \geq 0 \quad 0 < \beta < 1 \quad \gamma > 0$$

where Y is income, C is consumption, and I is investment.

7. Consider the following model of a market:

$$Q_{dt} = \alpha - \beta P_t \quad (\text{demand})$$

$$Q_{st} = -\gamma + \delta P_t^* \quad (\text{supply})$$

$$Q_{dt} = Q_{st}$$

$$\alpha > 0 \quad \beta > 0 \quad \gamma > 0 \quad \delta > 0$$

where P_t^* denotes the expected price in period t and sellers have an "adaptive" price expectation so that

$$P_t^* = P_{t-1}^* + \eta(P_{t-1} - P_{t-1}^*) \quad 0 < \eta < 1$$

- a. Show that the model can be represented by the first-order difference equation

$$P_{t+1} - \left(1 - \eta - \frac{\eta\delta}{\beta}\right)P_t = \frac{\eta(\alpha + \gamma)}{\beta}$$

Hint: Solve the supply equation for P_t^* and note that $Q_{st} = Q_{dt} = \alpha - \beta P_t$.

- b. Determine the nature of the solution sequence.

- c. Show that the solution sequence, if oscillatory, converges only if $1 - \frac{2}{\eta} < -\frac{\delta}{\beta}$.

8. In the Harrod model, graph $\{Y_t\}$ for each of the following cases and comment on the plausibility of each.

Case	α	β	Y_0
(a)	$\frac{1}{3}$	2	100
(b)	$-\frac{1}{3}$	2	100
(c)	-4	2	100
(d)	$\frac{1}{3}$	2	-100

Answers to Odd-Numbered Problems

1. If $\delta < 0$ and $\beta < 0$, the approach to equilibrium will not be oscillatory and $|\delta\beta| < 1$ (the stability condition) indicates that the demand curve must be flatter than the supply curve.

$$3. \quad \dot{Y}_t = \left(\frac{\delta\gamma}{\delta\gamma - \alpha} \right)^t \left(Y_0 + \frac{\beta}{\alpha} \right) - \frac{\beta}{\alpha} \quad \text{approaches } +\infty$$

$$S_t = \left(\frac{\delta\gamma}{\delta\gamma - \alpha} \right)^t (\alpha Y_0 + \beta) \quad \text{approaches } +\infty$$

$$I_t = \left(\frac{\delta\gamma}{\delta\gamma - \alpha} \right)^t \left(\frac{\alpha}{\delta} Y_0 + \frac{\beta}{\delta} \right) \quad \text{approaches } +\infty$$

$$\delta\gamma > \alpha$$

$$7. \quad b. \quad 0 < 1 - \eta - \frac{\eta\delta}{\beta} < 1 \quad \text{converges (nonoscillatory)}$$

$$-1 < 1 - \eta - \frac{\eta\delta}{\beta} < 0 \quad \text{converges (oscillatory)}$$

$$1 - \eta - \frac{\eta\delta}{\beta} < -1 \quad \text{diverges (oscillatory)}$$

$$1 - \eta - \frac{\eta\delta}{\beta} = -1 \quad \text{diverges (oscillatory)}$$

6.6 Linear Second-Order Difference Equations with Constant Coefficients

The method of solution for linear second-order difference equations with constant coefficients and the behavior of the solution sequence are discussed in this section. Special methods of solution are required when second-order difference equations are nonhomogeneous; these methods are considered and equilibrium and stability are defined.

The general second-order difference equation with constant coefficients can be written in the form

$$y_{x+2} + A_1 y_{x+1} + A_2 y_x = g(x)$$

Consider first the special case $g(x) = 0$:

$$y_{x+2} + A_1 y_{x+1} + A_2 y_x = 0$$

An equation whose constant term is zero is sometimes said to be *homogeneous*. Thus $y_{x+2} + A_1 y_{x+1} + A_2 y_x = 0$ is the general homogeneous second-order linear difference equation with constant coefficients. The definition of a homogeneous equation should not be confused with the definition of a homogeneous function; the two uses of the word homogeneous are not related.

In order to obtain the solution of

$$y_{x+2} + A_1 y_{x+1} + A_2 y_x = 0$$

form the auxiliary equation

$$m^2 + A_1 m + A_2 = 0$$

and solve for its roots, using the quadratic formula (if necessary):

$$m_1 = \frac{-A_1 + \sqrt{A_1^2 - 4A_2}}{2}$$

$$m_2 = \frac{-A_1 - \sqrt{A_1^2 - 4A_2}}{2}$$

These roots m_1 and m_2 may be real and unequal, real and equal, or complex (involving the square root of a negative number). The form of the solution of the equation $y_{x+2} + A_1 y_{x+1} + A_2 y_x = 0$ depends on the nature of the roots m_1 and m_2 , as summarized below. The general solution of a second-order difference equation includes two arbitrary constants and a particular solution is specified by two boundary conditions, that is, by two consecutive values of y .

Case 1. m_1 and m_2 are real and unequal ($m_1 \neq m_2$)

$$\text{solution: } y_x = C_1 m_1^x + C_2 m_2^x$$

Case 2. m_1 and m_2 are real and equal ($m_1 = m_2 = m$)

$$\text{solution: } y_x = C_1 m^x + C_2 x m^x$$

Case 3. m_1 and m_2 are complex ($m_1 = a + bi$, $m_2 = a - bi$ where $i = \sqrt{-1}$)

$$\text{solution: } y_x = r^x (C_1 \cos \theta x + C_2 \sin \theta x), \text{ where } r = \sqrt{a^2 + b^2} \text{ and } \theta \text{ is the angle having } \tan \theta = b/a; \text{ alternatively, } \theta \text{ is the angle having } \sin \theta = b/r, \cos \theta = a/r.$$

Note that a and b can be interchanged in these definitions, since the trigonometric functions are periodic. For convenience, the solution is usually written for θ in the first quadrant.

EXAMPLE

Obtain the general solution for the difference equation

$$y_{x+2} - 5y_{x+1} + 6y_x = 0$$

and the particular solution if $y_0 = 2$, $y_1 = 5$.

The auxiliary equation is

$$m^2 - 5m + 6 = 0$$

and

$$m_1 = \frac{5 + \sqrt{25 - 24}}{2} = 3$$

$$m_2 = \frac{5 - \sqrt{25 - 24}}{2} = 2$$

Thus the general solution is

$$y_x = C_1 3^x + C_2 2^x$$

If $y_0 = 2$, $y_1 = 5$,

$$2 = C_1 + C_2$$

$$5 = 3C_1 + 2C_2$$

$$C_2 = 1$$

$$C_1 = 1$$

and the particular solution is

$$y_x = 3^x + 2^x$$

EXAMPLE

Obtain the general solution for the difference equation

$$y_{x+2} - 4y_{x+1} + 4y_x = 0$$

and the particular solution if $y_0 = 1$, $y_1 = 6$.

The auxiliary equation is

$$m^2 - 4m + 4 = 0$$

and

$$m_1 = \frac{4 + \sqrt{16 - 16}}{2} = 2$$

$$m_2 = \frac{4 - \sqrt{16 - 16}}{2} = 2$$

Thus the general solution is

$$y_x = C_1 2^x + C_2 x 2^x$$

If $y_0 = 1$, $y_1 = 6$,

$$1 = C_1 + 0$$

$$6 = 2C_1 + 2C_2$$

$$C_1 = 1$$

$$C_2 = 2$$

and the particular solution is

$$y_x = 2^x + x 2^{x+1}$$

EXAMPLE

Obtain the general solution for the difference equation

$$y_{x+2} - y_{x+1} + \frac{1}{2}y_x = 0$$

and the particular solution if $y_0 = 3$, $y_1 = \frac{1}{2}$.

The auxiliary equation is

$$m^2 - m + \frac{1}{2} = 0$$

and

$$m_1 = \frac{1 + \sqrt{1 - 2}}{2} = \frac{1}{2} + \frac{1}{2}i$$

$$m_2 = \frac{1 - \sqrt{1 - 2}}{2} = \frac{1}{2} - \frac{1}{2}i$$

$$r = \sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2} = \frac{\sqrt{2}}{2}$$

$$\tan \theta = 1$$

$$\theta = \frac{1}{4}\pi$$

Thus the general solution is

$$y_x = \left(\frac{\sqrt{2}}{2}\right)^x (C_1 \cos \tfrac{1}{4}\pi x + C_2 \sin \tfrac{1}{4}\pi x)$$

If $y_0 = 3$, $y_1 = \frac{5}{2}$,

$$3 = (1)[(C_1)(1) + (C_2)(0)]$$

$$C_1 = 3$$

$$\frac{5}{2} = \frac{\sqrt{2}}{2} \left[3 \left(\frac{\sqrt{2}}{2} \right) + C_2 \left(\frac{\sqrt{2}}{2} \right) \right]$$

$$\frac{5}{2} = \frac{3}{2} + \frac{1}{2}C_2$$

$$C_2 = 2$$

and the particular solution is

$$y_x = \left(\frac{\sqrt{2}}{2}\right)^x (3 \cos \tfrac{1}{4}\pi x + 2 \sin \tfrac{1}{4}\pi x)$$

BEHAVIOR OF THE SOLUTION SEQUENCE

The behavior of the solution sequence depends on both the difference equation and the initial conditions; the roots of the auxiliary equation indicate the limiting behavior of the solution sequence as follows.

Case 1. Real roots, $m_1 \neq m_2$

If m_1 is the root with the greater absolute value, that is, $|m_1| > |m_2|$, then the limiting behavior of the solution sequence $\{C_1 m_1^x + C_2 m_2^x\}$ is the same as that of $\{C_1 m_1^x\}$, provided $C_1 \neq 0$.

This can be shown by writing

$$C_1 m_1^x + C_2 m_2^x = m_1^x [C_1 + C_2 (m_2/m_1)^x]$$

Since $-1 < m_2/m_1 < 1$, $(m_2/m_1)^x \rightarrow 0$ as $x \rightarrow \infty$ and

$$\frac{C_1 m_1^x}{C_1 m_1^x + C_2 m_2^x} = \frac{C_1}{C_1 + C_2 (m_2/m_1)^x} \rightarrow \frac{C_1}{C_1} = 1 \quad \text{as } x \rightarrow \infty$$

Thus the limiting behavior of $\{C_1 m_1^x + C_2 m_2^x\}$ is the same as the limiting behavior of $\{C_1 m_1^x\}$ if $|m_1| > |m_2|$. The limiting behavior of $\{C_1 m_1^x\}$ is discussed in the preceding sections concerning solution sequences of first-order difference equations. This limiting behavior can be summarized in the notation of the present section as follows.

If $|m_1| \leq 1$, the sequence converges

If $|m_1| > 1$, the sequence diverges

If $-1 < m_1 < 0$, the sequence is damped oscillatory

If $m < -1$, the sequence oscillates infinitely

If $C_1 = 0$, the solution sequence is $\{C_2 m_2^x\}$ and similar considerations are applicable to this sequence. Note that if $C_1 = 0$, $y_0 = C_2$ and $y_1 = C_2 m_2$. Thus $C_1 = 0$ implies boundary conditions which specify a particular solution.

Case 2. Real roots, $m_1 = m_2 = m$

The solution sequence is $\{(C_1 + C_2 x)m^x\}$, which diverges if $|m| > 1$ unless $C_1 = C_2 = 0$ and also diverges if $|m| = 1$ unless $C_2 = 0$. If $|m| < 1$, the sequence $\{xm^x\}$ converges to zero and $\{(C_1 + C_2 x)m^x\}$ also converges to zero. If m is negative, the sequence is oscillatory.

Case 3. Complex roots, $m_1 = a + bi$, $m_2 = a - bi$

The solution sequence is oscillatory; it converges to zero if $0 < \sqrt{a^2 + b^2} < 1$ and diverges if $\sqrt{a^2 + b^2} > 1$.

A complete classification of the solution sequences of the second-order linear homogeneous difference equation $y_{x+2} + A_1 y_{x+1} + A_2 y_x = 0$ is somewhat involved, since for exceptional initial values C_1 or C_2 or both are zero. There is one case in which the solution sequence for a homogeneous second-order linear difference equation converges to zero for every possible pair of initial values, as summarized in the following theorem.

THEOREM: If $\rho = \max(|m_1|, |m_2|)$, where m_1 and m_2 are the roots of the auxiliary equation of the homogeneous second-order difference equation,

$$y_{x+2} + A_1 y_{x+1} + A_2 y_x = 0$$

then $\rho < 1$ is a necessary and sufficient condition for the solution sequence $\{y_x\}$ to converge with limit 0 for all initial values y_0 and y_1 .

More specifically, the definition of ρ is as follows.

Case 1. m_1 and m_2 real and unequal

$$\rho = \max(|m_1|, |m_2|)$$

Case 2. m_1 and m_2 real and equal, $m_1 = m_2 = m$

$$\rho = |m|$$

Case 3. m_1 and m_2 complex, $m_1 = a + bi$, $m_2 = a - bi$

$$\rho = \sqrt{a^2 + b^2}$$

EXAMPLE

In the preceding example the general solution for the difference equation $y_{x+2} - 5y_{x+1} + 6y_x = 0$ is found to be

$$y_x = C_1 3^x + C_2 2^x$$

and, since $m_1 = 3$ and $m_2 = 2$, $\rho > 1$ and the solution sequence diverges, as can be seen from the first few values of the particular solution $y_x = 3^x + 2^x$: $y_0 = 2$, $y_1 = 5$, $y_2 = 13$, $y_3 = 35$, $y_4 = 97$, $y_5 = 275$, and so forth.

In another previous example, the general solution for the difference equation $y_{x+2} - 4y_{x+1} + 4y_x = 0$ is found to be

$$y_x = C_1 2^x + C_2 x 2^x$$

and, since $m_1 = m_2 = 2$, $\rho > 1$ and the solution sequence diverges, as can be seen from the first few values of the particular solution $y_x = 2^x + x 2^{x-1}$: $y_0 = 1$, $y_1 = 6$, $y_2 = 20$, $y_3 = 56$, $y_4 = 144$, $y_5 = 352$, and so forth.

In another previous example, the general solution for the difference equation $y_{x+2} - y_{x+1} + \frac{1}{2}y_x = 0$ is found to be

$$y_x = \left(\frac{\sqrt{2}}{2}\right)^x (C_1 \cos \frac{1}{4}\pi x + C_2 \sin \frac{1}{4}\pi x)$$

and $\rho = \sqrt{2}/2 < 1$, so the solution sequence is oscillatory and converges to zero, as can be seen from the first few values of the particular solution

$$y_x = \left(\frac{\sqrt{2}}{2}\right)^x (3 \cos \frac{1}{4}\pi x + 2 \sin \frac{1}{4}\pi x)$$

$y_0 = 3$, $y_1 = \frac{5}{2}$, $y_2 = 1$, $y_3 = -\frac{1}{4}$, $y_4 = -\frac{3}{4}$, $y_5 = -\frac{5}{8}$, and so forth.

NONHOMOGENEOUS SECOND-ORDER DIFFERENCE EQUATIONS

If the second-order linear homogeneous difference equation

$$y_{x+2} + A_1 y_{x+1} + A_2 y_x = 0$$

has the general solution y_x , then the second-order linear nonhomogeneous difference equation

$$y_{x+2} + A_1 y_{x+1} + A_2 y_x = g(x)$$

has the general solution $y_x + y_p$, where y_p is any particular solution of the nonhomogeneous equation.

Unfortunately, there is no general rule for determining y_p . The form of y_p depends on the form of $g(x)$; for simpler forms, y_p is similar to $g(x)$, but there are many special cases.

When $g(x)$ is a constant, y_p can be determined quite easily. Suppose $g(x)$ is a constant, K , that is,

$$y_{x+2} + A_1 y_{x+1} + A_2 y_x = K$$

First try a solution of the form: y_p equals a constant; say, $y_p = k$. If the solution $y_p = k$ satisfies the difference equation, then

$$k + A_1 k + A_2 k = K$$

and

$$k = \frac{K}{1 + A_1 + A_2}$$

Thus $y_p = K/(1 + A_1 + A_2)$ is a particular solution of the difference equation $y_{x+2} + A_1 y_{x+1} + A_2 y_x = K$, unless $1 + A_1 + A_2 = 0$.

If $1 + A_1 + A_2 = 0$, try a solution of the form $y_p = kx$. Then

$$k(x+2) + A_1 k(x+1) + A_2 kx = K$$

$$k(1 + A_1 + A_2)x + k(2 + A_1) = K$$

But $1 + A_1 + A_2 = 0$, and thus

$$k = \frac{K}{A_1 + 2}$$

and $y_p = K/(A_1 + 2)$, unless $1 + A_1 + A_2 = 0$ and $A_1 + 2 = 0$.

If both $1 + A_1 + A_2 = 0$ and $A_1 = -2$, that is, if $A_1 = -2$ and $A_2 = 1$, then try a solution of the form $y_p = kx^2$. It can be verified that, in this case, $y_p = (K/2)x^2$. Note that this solution is appropriate only for the difference equation $y_{x+2} - 2y_{x+1} + y_x = K$. Except for this one case, if $g(x)$ is a constant, the difference equation $y_{x+2} + A_1 y_{x+1} + A_2 y_x = g(x)$ has a particular solution, y_p , which is a constant.

In summary, when $g(x)$ is constant, the difference equation $y_{x+2} + A_1 y_{x+1} + A_2 y_x = g(x)$ has a particular solution as follows:

$$\begin{aligned} y_p &= \frac{K}{1 + A_1 + A_2} && \text{if } 1 + A_1 + A_2 \neq 0 \\ y_p &= \frac{K}{A_1 + 2} && \begin{aligned} &\text{if } 1 + A_1 + A_2 = 0 \\ &A_1 + 2 \neq 0 \end{aligned} \\ y_p &= \frac{K}{2} x^2 && \begin{aligned} &\text{if } 1 + A_1 + A_2 = 0 \\ &A_1 + 2 = 0 \\ &(\text{that is, if } A_1 = -2, A_2 = 1) \end{aligned} \end{aligned}$$

EXAMPLE

Find the general solution of the equation

$$y_{x+2} - 6y_{x+1} + 8 = 9$$

and a particular solution if $y_0 = 10$, $y_1 = 25$.

The auxiliary equation of the homogeneous equation is

$$m^2 - 6m + 8 = 0$$

$$(m - 4)(m - 2) = 0$$

$$m_1 = 4$$

$$m_2 = 2$$

The general solution of the homogeneous equation is thus

$$y_x = C_1 4^x + C_2 2^x$$

and the general solution of the nonhomogeneous equation is

$$y_x = C_1 4^x + C_2 2^x + y_p$$

where $y_p = K/(1 + A_1 + A_2) = 9/(1 - 6 + 8) = 3$ and thus

$$y_x = C_1 4^x + C_2 2^x + 3$$

If $y_0 = 10$, $y_1 = 25$,

$$C_1 + C_2 = 10 - 3 = 7$$

$$4C_1 + 2C_2 = 25 - 3 = 22$$

$$2C_1 = 8$$

$$C_1 = 4$$

$$C_2 = 3$$

and the particular solution is

$$y_x = 4^{x+1} + (3)2^x + 3$$

EXAMPLE

Find the general solution of the equation

$$y_{x+2} + 3y_{x+1} - 4 = 10$$

and a particular solution if $y_0 = 5$, $y_1 = 20$.

The auxiliary equation of the homogeneous equation is

$$m^2 + 3m - 4 = 0$$

$$(m + 4)(m - 1) = 0$$

$$m_1 = -4$$

$$m_2 = 1$$

The general solution of the homogeneous equation is thus

$$y_x = C_1 (-4)^x + C_2 (1)^x$$

and the general solution of the nonhomogeneous equation is

$$y_x = C_1 (-4)^x + C_2 + y_p$$

where $y_p = K/(A_1 + 2) = 10/5 = 2$, since $1 + A_1 + A_2 = 1 + 3 - 4 = 0$. Thus

$$y_x = C_1(-4)^x + C_2 + 2$$

If $y_0 = 5$, $y_1 = 20$,

$$C_1 + C_2 = 5 - 2 = 3$$

$$-4C_1 + C_2 = 20 - 2 = 18$$

$$-5C_1 = 15$$

$$C_1 = -3$$

$$C_2 = 6$$

and the particular solution is

$$y_x = -3(-4)^x + 8$$

Particular solutions of nonhomogeneous difference equations for which $g(x)$ is not constant, but is in fact a function of x , can be obtained similarly by examining trial solutions. However, this procedure may be quite tedious when $g(x)$ is not constant; these cases are not discussed here.

EQUILIBRIUM AND STABILITY

If a nonhomogeneous difference equation is of the form $y_{x+2} + A_1 y_{x+1} + A_2 y_x = K$, it can be shown that the solution sequence, of the corresponding homogeneous equation converges to zero if and only if $2 \pm A_1 > 1 - A_2 > 0$, which implies that $1 + A_1 + A_2 \neq 0$. Thus the solution sequence of the homogeneous equation converges to zero only when the nonhomogeneous equation has a particular solution $y_p = K/(1 + A_1 + A_2)$.

If the solution sequence of the corresponding homogeneous equation converges to zero, the solution of the nonhomogeneous equation $y_{x+2} + A_1 y_{x+1} + A_2 y_x = K$ converges to $K/(1 + A_1 + A_2)$; if the solution of the corresponding homogeneous equation diverges, the solution of the nonhomogeneous equation also diverges.

If the equation $y_{x+2} + A_1 y_{x+1} + A_2 y_x = K$ has a particular solution $y_p = K/(1 + A_1 + A_2)$, then y_p is an *equilibrium or stationary value* of y , denoted y^* .

An equilibrium value has the property that if two consecutive values of a solution of the equation $y_{x+2} + A_1 y_{x+1} + A_2 y_x = K$ are equal to y^* , then all succeeding values are equal to y^* . Substituting $y_x = y_{x+1} = y^* = K/(1 + A_1 + A_2)$ into $y_{x+2} + A_1 y_{x+1} + A_2 y_x = K$,

$$y_{x+2} + A_1 y^* + A_2 y^* = K$$

$$y_{x+2} = K - (A_1 + A_2)y^*$$

$$= K - \frac{K(A_1 + A_2)}{1 + A_1 + A_2}$$

$$= \frac{K}{1 + A_1 + A_2} = y^*$$

The equilibrium value is said to be *stable* or the difference equation $y_{x+2} + A_1 y_{x+1} + A_2 y_x = K$ is said to be *stable* if every solution of the difference equation converges to y^* for every possible set of initial conditions y_0 and y_1 , that is, if

$$\lim y_x = y^* \quad \text{for all } y_0 \text{ and } y_1$$

It can be shown that a necessary and sufficient condition for the equilibrium value $y^* = K/(1 + A_1 + A_2)$ to be stable is $\rho < 1$, where, as above, $\rho = \max(|m_1|, |m_2|)$ and m_1 and m_2 are the roots of the auxiliary equation $m^2 + A_1 m + A_2 = 0$.

It can be shown that $\rho < 1$ if and only if $2 \pm A_1 > 1 - A_2 > 0$. Thus these inequalities are necessary and sufficient for stability of the equilibrium value y^* and $y^* = K/(1 + A_1 + A_2)$ is a stable equilibrium for the difference equation $y_{x+2} + A_1 y_{x+1} + A_2 y_x = K$ if and only if $2 \pm A_1 > 1 - A_2 > 0$.

PROBLEMS

Obtain the general solution for each of the following difference equations.

1. $y_{x+2} + 2y_{x+1} + y_x = 0$
2. $y_{x+2} - y_x = 0$
3. $2y_{x+2} - 5y_{x+1} + 2y_x = 0$
4. $3y_{x+2} - 6y_{x+1} + 4y_x = 0$
5. Find the general solution for the difference equation $y_{x+2} + 2y_x = 0$ and the particular solution if $y_0 = 1$, $y_1 = \sqrt{2}$.
6. Find the general solution for the difference equation $y_{x+2} + 3y_{x+1} + 3y_x = 0$ and the particular solution if $y_0 = 3$, $y_1 = 0$.

For each of the following difference equations, obtain the general solution and the particular solution for the specified initial values.

7. $y_{x+2} + 4y_{x+1} + 8y_x = 26$, $y_0 = 6$, $y_1 = 3$
8. $y_{x+2} + 8y_{x+1} + 16y_x = 25$, $y_0 = 0$, $y_1 = 4$
9. $y_{x+2} - 8y_{x+1} - 9y_x = 24$, $y_0 = 2$, $y_1 = 0$
10. $3y_{x+2} - 10y_{x+1} + 3y_x = 8$, $y_0 = 5$, $y_1 = 3$

Solve each of the following difference equations, determine the behavior of the solution sequence, and calculate the first few values of the solution sequence.

11. $y_{x+2} - 3y_{x+1} + 3y_x = 5$, $y_0 = 5$, $y_1 = 8$
12. $9y_{x+2} - 6y_{x+1} + y_x = 16$, $y_0 = 0$, $y_1 = 3$
13. $6y_{x+2} + 5y_{x+1} - y_x = 20$, $y_0 = 3$, $y_1 = 8$
14. $4y_{x+2} - y_x = 15$, $y_0 = 15$, $y_1 = 10$
15. $8y_{x+2} - 6y_{x+1} + y_x = 9$, $y_0 = 10$, $y_1 = 15$
16. $y_{x+2} - 4y_{x+1} + 4y_x = 1$, $y_0 = 0$, $y_1 = 1$
17. $y_{x+2} - 5y_{x+1} + 6y_x = 4$, $y_0 = 0$, $y_1 = 1$
18. $y_{x+2} - 7y_{x+1} + 12y_x = 2$, $y_0 = 0$, $y_1 = 1$
19. $y_{x+2} - 2y_{x+1} + 2y_x = 3$, $y_0 = 5$, $y_1 = 6$
20. $y_{x+2} - 4y_x = 9$, $y_0 = 0$, $y_1 = 1$
21. $12y_{x+2} - 7y_{x+1} + y_x = 18$, $y_0 = 0$, $y_1 = 3$
22. $3y_{x+2} + 5y_{x+1} + 2y_x = 4$, $y_0 = 0$, $y_1 = 1$

Answers to Odd-Numbered Problems

1. $y_x = (-1)^x (C_1 + C_2 x)$
3. $y_x = C_1 \left(\frac{1}{2}\right)^x + C_2 (2)^x$
5. $y_x = (\sqrt{2})^x (\cos \frac{1}{2}\pi x + \sin \frac{1}{2}\pi x)$
7. $y_x = (2\sqrt{2})^x (4 \cos \frac{1}{4}\pi x - \frac{7}{2} \sin \frac{1}{4}\pi x) + 2$
9. $y_x = 3(-1)^x + \frac{1}{2}(9)^x - \frac{3}{2}$
11. $y_x = 2(\sqrt{3})^x (\sin \frac{1}{3}\pi x) + 5$, divergent (oscillatory), $y_0 = 5$, $y_1 = 8$, $y_2 = 5 + 3\sqrt{3}$, $y_3 = 5$, $y_4 = 5 - 9\sqrt{3}$, $y_5 = -22$
13. $y_x = -5(-1)^x + \left(\frac{1}{6}\right)^{x-1} + 2$, divergent (oscillatory), $y_0 = 3$, $y_1 = 8$, $y_2 = -2\frac{5}{6}$, $y_3 = 7\frac{1}{36}$, $y_4 = -2\frac{215}{116}$, $y_5 = 7\frac{1}{1296}$
15. $y_x = 6\left(\frac{1}{4}\right)^x + \left(\frac{1}{2}\right)^x + 3$, convergent, $y_0 = 10$, $y_1 = 5$, $y_2 = 3\frac{5}{8}$, $y_3 = 3\frac{7}{32}$, $y_4 = 3\frac{11}{128}$, $y_5 = 3\frac{19}{512}$
17. $y_x = 3^{x+1} - 5(2)^x + 2$, divergent, $y_0 = 0$, $y_1 = 1$, $y_2 = 9$, $y_3 = 43$, $y_4 = 165$, $y_5 = 571$

19. $y_x = (\sqrt{2})^x (2 \cos \frac{1}{4}\pi x + \sin \frac{1}{4}\pi x) + 3$, divergent (oscillatory), $y_0 = 5$, $y_1 = 6$, $y_2 = 5$, $y_3 = 1$, $y_4 = -5$, $y_5 = -9$
21. $y_x = -12(\frac{1}{4})^x + 9(\frac{1}{3})^x + 3$, convergent, $y_0 = 0$, $y_1 = 3$, $y_2 = 3\frac{1}{4}$, $y_3 = 3\frac{7}{48}$, $y_4 = 3\frac{37}{576}$, $y_5 = 3\frac{175}{6912}$

6.7 Applications of Second-Order Difference Equations in Economic Models

Second-order difference equations are used in many economic models to indicate the dependence of the value of a variable on its value for two preceding periods. Second-order difference equations may be associated with various types of relationships among variables in the model; some of these relationships are illustrated in Samuelson's interaction model and Metzler's inventory model, which are discussed in this section.

SAMUELSON'S INTERACTION MODEL

The following model was proposed by P. A. Samuelson for national income analysis.

$$Y_t = C_t + I_t + G_t$$

$$C_t = \alpha Y_{t-1}$$

$$I_t = \beta [C_t - C_{t-1}]$$

$$Y_0 = Y_0$$

$$Y_1 = Y_1$$

$$\alpha > 0 \quad \beta > 0$$

where Y is national income, C is consumption, I is investment, and G is government expenditure. G is assumed constant (exogenous) from period to period and, for convenience, the units are assumed to be such that $G = 1$. Note that investment in any period is equal to a constant times the increase in consumption of that period over the preceding period; this is referred to as the acceleration principle.

Substituting the second and third equations into the first equation,

$$Y_t = \alpha Y_{t-1} + \beta(\alpha Y_{t-1} - \alpha Y_{t-2}) + 1$$

$$Y_t - \alpha(1 + \beta)Y_{t-1} + \alpha\beta Y_{t-2} - 1 = 0$$

The three types of general solution of the homogeneous equation, depending on the nature of the roots of the auxiliary equation, are as follows:

(a) $Y_t = C_1 m_1^t + C_2 m_2^t$, where

$$m_1 = \frac{\alpha(1 + \beta) + \sqrt{\alpha^2(1 + \beta)^2 - 4\alpha\beta}}{2}$$

$$m_2 = \frac{\alpha(1 + \beta) - \sqrt{\alpha^2(1 + \beta)^2 - 4\alpha\beta}}{2}$$

(b) $Y_t = C_1 m^t + C_2 t m^t$, where

$$m = \frac{\alpha}{2}(1 + \beta)$$

(c) $Y_t = r^t(C_1 \cos \theta t + C_2 \sin \theta t)$, where

$$r = \sqrt{\frac{\alpha^2(1 + \beta^2)}{4} + \frac{4\alpha\beta - \alpha^2(1 + \beta^2)}{4}} = \sqrt{\alpha\beta}$$

$$\sin \theta = \frac{\alpha(1 + \beta)}{2\sqrt{\alpha\beta}}$$

$$\cos \theta = \sqrt{\frac{4\alpha\beta - \alpha^2(1 + \beta^2)}{4\alpha\beta}} = \sqrt{1 - \frac{\alpha(1 + \beta)}{4\beta}}$$

A particular solution of the nonhomogeneous equation

$$Y_t - \alpha(1 + \beta)Y_{t-1} + \alpha\beta Y_{t-2} - 1 = 0$$

is given by

$$\begin{aligned} y_p &= \frac{1}{1 - \alpha(\alpha + \beta) + \alpha\beta} \\ &= \frac{1}{1 - \alpha} \end{aligned}$$

assuming that $1 - \alpha \neq 0$. Since α is the marginal propensity to consume, $\alpha < 1$ and $1 - \alpha \neq 0$.

Thus the solution of the nonhomogeneous equation has the form

$$(a) \quad Y_t = C_1 m_1^t + C_2 m_2^t + \frac{1}{1 - \alpha}$$

$$(b) \quad Y_t = C_1 m^t + C_2 t m^t + \frac{1}{1 - \alpha}$$

$$(c) \quad Y_t = r^t(C_1 \cos \theta t + C_2 \sin \theta t) + \frac{1}{1 - \alpha}$$

where m_1 , m_2 , m , r , and θ are defined as above and the arbitrary constants can be determined from the initial conditions Y_0 and Y_1 . Whenever the sum of the first two terms of the solution (in any of the three possible forms) tends to zero, $y^* = y_p = 1/(1 - \alpha)$ is an equilibrium solution.

The stability conditions $1 + A_1 + A_2 > 0$, $1 - A_1 + A_2 > 0$ and $1 - A_2 > 0$ become

$$1 - \alpha(1 + \beta) + \alpha\beta > 0 \quad \text{or} \quad 1 - \alpha > 0$$

$$1 + \alpha(1 + \beta) + \alpha\beta > 0 \quad \text{or} \quad 1 + \alpha + 2\alpha\beta > 0$$

$$1 - \alpha\beta > 0$$

Since $\alpha > 0$, $\beta > 0$, the second condition is satisfied and the necessary and sufficient conditions for y^* to be a stable equilibrium value are

$$\alpha < 1$$

$$\alpha\beta < 1$$

That is, both the marginal propensity to consume (one year's consumption with respect to the previous year's income) and its product with the accelerator parameter must be less than 1 for the sequence of income values to converge to y^* for all possible initial conditions. The convergence to y^* is oscillatory if the roots of the auxiliary equation are complex numbers. This is the case if $A_1^2 - 4A_2$ is negative, that is, if $\alpha^2(1 + \beta)^2 - 4\alpha\beta < 0$.

METZLER'S INVENTORY MODEL

The following model was proposed by L. A. Metzler for analyzing inventory cycles.

$$y_t = u_t + s_t + v_0$$

$$u_t = \beta y_{t-1}$$

$$s_t = \beta(y_{t-1} - y_{t-2})$$

$$0 < \beta < 1$$

where y_t is total income produced in period t , u_t is consumers' goods produced for sale in period t , s_t is consumers' goods produced for inventories in period t , and v_0 is the constant noninduced net investment in each period.

The total income produced in any period is equal to the total production of consumers' goods plus net investment. Sales in any period are a constant proportion of the income in the preceding period. Production for inventory is equal to the difference between actual and anticipated sales of the preceding period; that is, there is an attempt to keep inventory at a constant level. It is assumed that inventories are sufficient to meet differences between production and consumer demand.

Substituting the second and third equations into the first equation,

$$y_t = \beta y_{t-1} + \beta(y_{t-1} - y_{t-2}) + v_0$$

$$y_t - 2\beta y_{t-1} + \beta y_{t-2} = v_0$$

equivalently, so the t -values can begin at $t = 0$,

$$y_{t+2} - 2\beta y_{t+1} + \beta y_t = v_0$$

The auxiliary equation of the homogeneous difference equation corresponding to the equation $y_{t+2} - 2\beta y_{t+1} + \beta y_t = v_0$ is

$$m^2 - 2\beta m + \beta = 0$$

$$m = \frac{2\beta \pm \sqrt{4\beta^2 - 4\beta}}{2} = \beta \pm \sqrt{\beta^2 - \beta}$$

Since $0 < \beta < 1$, $\beta^2 - \beta < 0$ and the roots are complex numbers

$$m = \beta \pm i\sqrt{\beta(1-\beta)}$$

Thus

$$r = \sqrt{\beta^2 + \beta(1-\beta)} = \sqrt{\beta}$$

$$\cos \theta = \frac{\beta}{\sqrt{\beta}} = \sqrt{\beta}$$

$$\sin \theta = \frac{\sqrt{\beta(1-\beta)}}{\sqrt{\beta}} = \sqrt{1-\beta}$$

The general solution of the homogeneous equation

$$y_{t+2} - 2\beta y_{t+1} + \beta y_t = 0$$

is thus given by

$$y_t = (\sqrt{\beta})^t (C_1 \cos \theta t + C_2 \sin \theta t)$$

A particular solution of the nonhomogeneous equation

$$y_{t+2} - 2\beta y_{t+1} + \beta y_t = v_0$$

is given by

$$y_p = \frac{v_0}{1 - \beta}$$

since $0 < \beta < 1$ and thus $1 - \beta \neq 0$. The general solution of the nonhomogeneous equation thus has the form

$$y_t = (\sqrt{\beta})^t (C_1 \cos \theta t + C_2 \sin \theta t) + \frac{v_0}{1 - \beta}$$

The term involving sines and cosines results in cyclical fluctuations because of its oscillation between positive and negative values. These fluctuations are damped by the factor $(\sqrt{\beta})^t$ since $0 < \beta < 1$. Thus

$$(\sqrt{\beta})^t (C_1 \cos \theta t + C_2 \sin \theta t) \rightarrow 0 \quad \text{as } t \rightarrow \infty$$

and

$$y_t \rightarrow \frac{v_0}{1 - \beta} \quad \text{as } t \rightarrow \infty$$

as an equilibrium value.

PROBLEMS

1. For the Samuelson model, verify the expressions for m_1 , m_2 , m , r , and θ given in the text.
2. For the Samuelson model, show that the roots of the auxiliary equation are complex if and only if

$$\alpha^2(1 + \beta)^2 - 4\alpha\beta < 0.$$

3. For the Samuelson model, obtain the solution and determine the behavior of the solution sequence if a. $\alpha = 0.8$, $\beta = 0.3$ and b. $\alpha = 0.5$, $\beta = 0.1$.
4. For the price-adjustment equation

$$p_{t+2} = \beta r_0 + \beta \alpha (p_{t+1} - p_t)$$

where p is price and β , α , and r_0 are positive constants, determine the limiting behavior of the solution if a. $\beta\alpha = 1$, b. $\beta\alpha = \frac{1}{2}$, c. $\beta\alpha = 2$, d. $\beta\alpha = \frac{3}{4}$.

5. Consider the following difference equation discussed by Baumol:

$$Y_{t+2} - \frac{C}{s} Y_{t+1} + \frac{C}{s} Y_t = 0$$

where Y is warranted income and C and s are positive constants. Show that for $C \geq 4s$ this equation gives an explosive (divergent) solution for warranted income, but for $C < 4s$ warranted income is cyclic.

6. For the Metzler model, note that (total) inventory, denoted by i , is given by

$$i_t = i_{t-1} + s_t + u_t - y_t$$

Obtain an expression for change in inventory, $i_t - i_{t-1}$, in terms of y_t and determine the behavior of this solution sequence.

Answers to Odd-Numbered Problems

3. a. $y_t = C_1(2)^t + C_2(\frac{5}{2})^t + 5$, divergent
b. $y_t = (\frac{1}{2}\sqrt{2})^t (C_1 \cos \frac{1}{4}\pi t + C_2 \sin \frac{1}{4}\pi t) + 2$, convergent (oscillatory) to 2

Matrix Algebra

7

7.1 Introduction

In many economic analyses, variables are assumed to be related by sets of linear equations. Matrix algebra provides a clear and concise notation for the formulation and solution of such problems, many of which would be almost impossibly complicated in conventional algebraic notation.

In this chapter, vectors, matrices, and operations involving them are defined. Special types of matrices, the transpose of a matrix, partitioned matrices, and the determinant of a matrix are considered. Linear dependence of a set of vectors and the rank and inverse of a matrix are also discussed and are applied to the solution of simultaneous equations. Vector differentiation is defined and illustrated. In Chapter 8 further applications of matrix algebra are considered.

7.2 Definition of a Matrix

In this section matrices are defined and illustrated. A *matrix* is a rectangular array of numbers written in the form

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \quad \text{or} \quad \mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

The letters a_{ij} stand for real numbers, the *elements* of the matrix. Note that a_{ij} is the element in the i th row and the j th column of the matrix \mathbf{A} ; the matrix \mathbf{A} is sometimes denoted by (a_{ij}) or by $\{a_{ij}\}$. A matrix that has m rows and n columns is said to be an $m \times n$ (read “ m by n ”) matrix or a matrix of *order* $m \times n$. If $m = n$, the matrix is said to be *square*. Especially when various operations are being performed on them, the order of matrices is frequently denoted by subscripts, for example, $\mathbf{A}_{m \times n}$ or $(a_{ij})_{m \times n}$.

EXAMPLE

$$\begin{bmatrix} 1 & 0 & -2 & 6 \\ 4 & 8 & 3 & -9 \end{bmatrix} \quad \text{is a } 2 \times 4 \text{ matrix}$$

$$\begin{bmatrix} 6 & 6 & 3 \\ 3 & 8 & -2 \\ -1 & 0 & 0 \end{bmatrix} \quad \text{is a } 3 \times 3 \text{ (square) matrix}$$

$$\begin{bmatrix} 5 & -8 & -2 \\ 12 & 10 & -1 \\ 13 & 9 & -3 \\ 2 & 7 & 6 \\ 6 & 4 & 10 \end{bmatrix} \quad \text{is a } 5 \times 3 \text{ matrix}$$

$$\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \quad \text{is a } 2 \times 2 \text{ (square) matrix}$$

Two matrices of the same order are said to be *equal* if and only if all the corresponding elements are equal, that is, if the matrices are identical.

EXAMPLE

If

$$A = \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}$$

$$B = \begin{bmatrix} 2 & -2 & 2 \\ -2 & 2 & -2 \end{bmatrix}$$

$$C = \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix}$$

$$D = \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}$$

$A = D$, but $A \neq B$, $A \neq C$, $B \neq C$, $B \neq D$, and $C \neq D$.

DEFINITION OF A VECTOR

A matrix that consists of a single column, that is, an $m \times 1$ matrix, is said to be a *column vector* and is written

$$\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{pmatrix} \quad \text{or} \quad \mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix}$$

The letters u_i stand for real numbers, the *components* of the vector; u_i is the i th component of the vector \mathbf{u} . A column vector that has m rows is said to be an *m-component* or an *m-dimensional* vector.

Similarly, a matrix that contains only a single row, that is, a $1 \times n$ matrix, is said to be a *row vector* and is written

$$\mathbf{v} = (v_1, v_2, \dots, v_n) \quad \text{or} \quad \mathbf{v} = [v_1, v_2, \dots, v_n]$$

The letters v_j stand for real numbers, the *components* of the vector; v_j is the j th component of the vector \mathbf{v} . A row vector that has n columns is said to be an n -*component* or an n -*dimensional vector*.

EXAMPLE

$$\begin{bmatrix} -1 \\ -1 \end{bmatrix} \quad \text{is a } 2 \times 1 \text{ matrix, a 2-dimensional column vector}$$

$$\begin{bmatrix} 0 \\ 0 \\ 3 \\ 0 \\ 2 \end{bmatrix} \quad \text{is a } 5 \times 1 \text{ matrix, a 5-dimensional column vector}$$

$$[0, 3, 0] \quad \text{is a } 1 \times 3 \text{ matrix, a 3-dimensional row vector}$$

$$[-1, -1, 5, -1] \quad \text{is a } 1 \times 4 \text{ matrix, a 4-dimensional row vector}$$

Two row vectors having the same number of rows or two column vectors having the same number of columns are said to be *equal* if and only if all the corresponding elements are equal, that is, if the vectors are identical.

EXAMPLE

$$\mathbf{u} = [1, 3]$$

$$\mathbf{v} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

$$\mathbf{w} = [1, 3]$$

$$\mathbf{x} = [3, 1]$$

$$\mathbf{u} = \mathbf{w}, \text{ but } \mathbf{u} \neq \mathbf{v}, \mathbf{u} \neq \mathbf{x}, \text{ and } \mathbf{v} \neq \mathbf{w}, \mathbf{v} \neq \mathbf{x}, \text{ and } \mathbf{w} \neq \mathbf{x}.$$

NOTE: Frequently it is useful to regard a matrix as being composed of a series of row or column vectors; for example, the matrix

$$\begin{bmatrix} 1 & -6 \\ 2 & 2 \\ -3 & 4 \end{bmatrix}$$

can be regarded as consisting of the two column vectors

$$\begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -6 \\ 2 \\ 4 \end{bmatrix}$$

or, alternatively, of the three row vectors

$$[1, -6] \quad [2, 2] \quad [-3, 4].$$

7.3 Matrix Operations

Operations analogous to the operations of addition, subtraction, multiplication, and division of real numbers can be defined for matrices. Since a matrix is an array of numbers, rather than a single number, some of the properties of operations for real numbers do not hold for the analogous matrix operations; specific examples are noted in following sections. In this section addition and subtraction of matrices, multiplication of a matrix by a scalar, and multiplication of matrices are defined and illustrated.

ADDITION AND SUBTRACTION OF MATRICES

Matrices can be added or subtracted if and only if they are of the same order. The sum or difference of two $m \times n$ matrices is another $m \times n$ matrix whose elements are the sums or differences of the corresponding elements in the component matrices; thus, if

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix}$$

then

$$\mathbf{A} \pm \mathbf{B} = \mathbf{C}$$

where

$$\mathbf{C} = \begin{bmatrix} a_{11} \pm b_{11} & \cdots & a_{1n} \pm b_{1n} \\ \vdots & & \vdots \\ a_{m1} \pm b_{m1} & \cdots & a_{mn} \pm b_{mn} \end{bmatrix} = \begin{bmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & & \vdots \\ c_{m1} & \cdots & c_{mn} \end{bmatrix}$$

that is, $(a_{ij}) + (b_{ij}) = (c_{ij})$, where $c_{ij} = a_{ij} + b_{ij}$ for all i and j .

EXAMPLE

$$(a) \begin{bmatrix} 3 & 2 & -4 \\ 5 & 6 & 8 \\ 3 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 3 & 8 \\ -5 & -6 & 2 \\ 0 & 0 & -4 \end{bmatrix} = \begin{bmatrix} 3 & 5 & 4 \\ 0 & 0 & 10 \\ 3 & 0 & -4 \end{bmatrix}$$

$$(b) \begin{bmatrix} 3 & -10 \\ -11 & 25 \end{bmatrix} - \begin{bmatrix} -6 & -4 \\ 22 & -21 \end{bmatrix} = \begin{bmatrix} 9 & -6 \\ -33 & 46 \end{bmatrix}$$

$$(c) [4, 6, 12] + [-3, 2, -12] = [1, 8, 0]$$

$$(d) \begin{bmatrix} 1 \\ 1 \\ 2 \\ 4 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 2 \\ 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$(e) \begin{bmatrix} 2 & 3 \\ 6 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 6 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ -1 & 2 \end{bmatrix}$$

$$(f) \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix} - \begin{bmatrix} 6 \\ 8 \\ 10 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ -4 \\ -5 \end{bmatrix}$$

$$(g) [2, 1] + [3, 4] + [6, 7] - [11, 12] = [0, 0]$$

$$(h) \begin{bmatrix} 1 & 4 \\ 2 & 6 \\ 3 & 8 \end{bmatrix} - \begin{bmatrix} 0 & 5 \\ 7 & 9 \\ 11 & 12 \end{bmatrix} - \begin{bmatrix} 10 & 13 \\ 20 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -9 & -14 \\ -25 & -3 \\ -9 & -4 \end{bmatrix}$$

PROBLEMS

Obtain the matrix resulting from each of the following operations.

$$1. \begin{bmatrix} 2 & -3 & 6 \\ 5 & 4 & 5 \\ 0 & -1 & -9 \end{bmatrix} - \begin{bmatrix} 1 & -3 & 4 \\ 0 & -2 & 5 \\ 1 & 0 & -1 \end{bmatrix}$$

$$2. \begin{bmatrix} 6 & -1 & 0 \\ 4 & 2 & 1 \end{bmatrix} + \begin{bmatrix} 5 & 0 & 2 \\ 0 & -1 & 3 \end{bmatrix} + \begin{bmatrix} -2 & -1 & -3 \\ -4 & 1 & -1 \end{bmatrix}$$

$$3. \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix} - \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix} + \begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix}$$

$$4. \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} + \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} - \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$$

$$5. [1, 3, -1, 2] - [0, 1, -2, 3]$$

$$6. [-1, 2] - [3, 4] + [1, -2] - [6, 5]$$

Answers to Odd-Numbered Problems

$$1. \begin{bmatrix} 1 & 0 & 2 \\ 5 & 6 & 0 \\ -1 & -1 & -8 \end{bmatrix} \quad 3. \begin{bmatrix} 2 \\ 4 \\ 0 \end{bmatrix}$$

$$5. [1, 2, 1, -1]$$

MULTIPLICATION OF A MATRIX BY A SCALAR

A real number (a 1×1 matrix) is referred to as a *scalar* when it occurs in operations involving matrices. When a matrix is multiplied by scalar, every element in the matrix is multiplied by that scalar (number); thus, if

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \quad \text{and } k \text{ is any scalar (constant)}$$

then

$$k \times \mathbf{A}_{m \times n} = k\mathbf{A}_{m \times n} = \begin{bmatrix} ka_{11} & \cdots & ka_{1n} \\ \vdots & & \vdots \\ ka_{m1} & \cdots & ka_{mn} \end{bmatrix} = k(a_{ij})_{m \times n} = (ka_{ij})_{m \times n}$$

EXAMPLE

$$(a) 3 \begin{bmatrix} 4 & -3 \\ 8 & -2 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 12 & -9 \\ 24 & -6 \\ -3 & 0 \end{bmatrix}$$

$$(b) \ 5 \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -5 \\ 0 \end{bmatrix}$$

$$(c) \ -1[6, -2, -3] = [-6, 2, 3]$$

$$(d) \ a \begin{bmatrix} 5 & 6 & 2 & 4 \\ -3 & -1 & 0 & -6 \end{bmatrix} = \begin{bmatrix} 5a & 6a & 2a & 4a \\ -3a & -a & 0 & -6a \end{bmatrix}$$

$$(e) \ -b \begin{bmatrix} 0 \\ -2 \\ 3 \\ -1 \\ 5 \end{bmatrix} = \begin{bmatrix} 0 \\ 2b \\ -3b \\ b \\ -5b \end{bmatrix}$$

$$(f) \ c[0, 0, 0, 0, 16] = [0, 0, 0, 0, 16c]$$

MULTIPLICATION OF MATRICES

Two matrices can be multiplied if and only if the number of columns in one matrix is equal to the number of rows in the other matrix. In particular, the matrix product \mathbf{AB} is defined if and only if the number of columns in \mathbf{A} is the same as the number of rows in \mathbf{B} , in this case the matrices \mathbf{A} and \mathbf{B} are said to be *conformable* for multiplication and the product matrix has the same number of rows as \mathbf{A} and the same number of columns as \mathbf{B} . Thus an $m \times n$ matrix may be multiplied times an $n \times p$ matrix to obtain an $m \times p$ matrix.

DEFINITION: When a $1 \times n$ row vector multiplies an $n \times 1$ column vector, the result is a scalar, the *inner product* of the two vectors; the value of the inner product is the sum of products of the components of the vectors. Thus if

$$\mathbf{u} = [u_1, \dots, u_n] \quad \text{and} \quad \mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

then $\mathbf{u}_{1 \times n} \mathbf{v}_{n \times 1} = w$ (a scalar), where $w = u_1 v_1 + u_2 v_2 + \dots + u_n v_n = \sum_{i=1}^n u_i v_i$.

When two matrices are multiplied, the element in the i th row and j th column of the product matrix is the inner product of the i th row vector of the first matrix with the j th column vector of the second matrix. Thus the product of two matrices may be written as a matrix of their inner products: If $\mathbf{A} = (a_{ij})_{m \times n}$ and $\mathbf{B} = (b_{jk})_{n \times p}$, then $\mathbf{AB} = \mathbf{C}$, where

$$\mathbf{C} = (c_{ik})_{m \times p} = \begin{bmatrix} \sum_{j=1}^n a_{1j} b_{j1} & \cdots & \sum_{j=1}^n a_{1j} b_{jm} \\ \vdots & & \vdots \\ \sum_{j=1}^n a_{mj} b_{j1} & \cdots & \sum_{j=1}^n a_{mj} b_{jn} \end{bmatrix}$$

that is, $c_{ik} = \sum_{j=1}^n a_{ij} b_{jk}$.

EXAMPLE

$$(a) \begin{bmatrix} 1 & 3 & -1 \\ 2 & 0 & 0 \\ 0 & -1 & 6 \end{bmatrix}_{3 \times 3} \begin{bmatrix} 1 & 0 \\ -1 & 2 \\ 1 & 3 \end{bmatrix}_{3 \times 2} = \begin{bmatrix} -3 & 3 \\ 2 & 0 \\ 7 & 16 \end{bmatrix}_{3 \times 2}$$

where

$$(1)(1) + (3)(-1) + (-1)(1) = -3$$

$$(2)(1) + (0)(-1) + (0)(1) = 2$$

$$(0)(1) + (-1)(-1) + (6)(1) = 7$$

$$(1)(0) + (3)(2) + (-1)(3) = 3$$

$$(2)(0) + (0)(2) + (0)(3) = 0$$

$$(0)(0) + (-1)(2) + (6)(3) = 16$$

$$(b) [-1, 0, 6, 3, 2]_{1 \times 5} \begin{bmatrix} 3 & 0 \\ 4 & 0 \\ -2 & 3 \\ 1 & 8 \\ 0 & -2 \end{bmatrix}_{5 \times 2} = [-12, 38]_{1 \times 2}$$

where

$$(-1)(3) + (0)(4) + (6)(-2) + (3)(1) + (2)(0) = -12$$

$$(-1)(0) + (0)(0) + (6)(3) + (3)(8) + (2)(-2) = 38$$

In matrix multiplication, the sequence in which multiplication is performed is very important. If \mathbf{A} is $m \times n$ and \mathbf{B} is $n \times m$, then it is possible to obtain both of the product matrices \mathbf{AB} and \mathbf{BA} ; however, in general, $\mathbf{AB} \neq \mathbf{BA}$. In the matrix product \mathbf{AB} , \mathbf{A} is said to *premultiply* \mathbf{B} or, alternatively, \mathbf{B} is said to *postmultiply* \mathbf{A} . Since, in general, premultiplication and postmultiplication give different results, even when both are defined, care must be taken to preserve the appropriate sequence in all matrix multiplications. Note that this precaution is not necessary in the multiplication of numbers.

EXAMPLE

(a) If

$$\mathbf{A} = \begin{bmatrix} 4 & 6 & -1 & 3 \\ 0 & -1 & 2 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 1 & 2 \\ -1 & 1 \\ 1 & 6 \\ 2 & 3 \end{bmatrix}$$

then

$$\mathbf{AB} = \begin{bmatrix} 4 & 6 & -1 & 3 \\ 0 & -1 & 2 & 1 \end{bmatrix}_{2 \times 4} \begin{bmatrix} 1 & 2 \\ -1 & 1 \\ 1 & 6 \\ 2 & 3 \end{bmatrix}_{4 \times 2} = \begin{bmatrix} 3 & 17 \\ 5 & 14 \end{bmatrix}_{2 \times 2}$$

$$\mathbf{BA} = \begin{bmatrix} 1 & 2 \\ -1 & 1 \\ 1 & 6 \\ 2 & 3 \end{bmatrix}_{4 \times 2} \begin{bmatrix} 4 & 6 & -1 & 3 \\ 0 & -1 & 2 & 1 \end{bmatrix}_{2 \times 4} = \begin{bmatrix} 4 & 4 & 3 & 5 \\ -4 & -7 & 3 & -2 \\ 4 & 0 & 11 & 9 \\ 8 & 9 & 4 & 9 \end{bmatrix}_{4 \times 4}$$

(b) If

$$\mathbf{A} = \begin{bmatrix} 5 & -6 \\ -1 & 0 \\ 0 & 3 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} -1 & 8 & -3 \\ 0 & 10 & -4 \end{bmatrix}$$

then

$$\mathbf{AB} = \begin{bmatrix} 5 & -6 \\ -1 & 0 \\ 0 & 3 \end{bmatrix}_{3 \times 2} \begin{bmatrix} -1 & 8 & -3 \\ 0 & 10 & -4 \end{bmatrix}_{2 \times 3} = \begin{bmatrix} -5 & -20 & 9 \\ 1 & -8 & 3 \\ 0 & 30 & -12 \end{bmatrix}_{3 \times 3}$$

$$\mathbf{BA} = \begin{bmatrix} -1 & 8 & -3 \\ 0 & 10 & -4 \end{bmatrix}_{2 \times 3} \begin{bmatrix} 5 & -6 \\ -1 & 0 \\ 0 & 3 \end{bmatrix}_{3 \times 2} = \begin{bmatrix} -13 & -3 \\ -10 & -12 \end{bmatrix}_{2 \times 2}$$

(c) If

$$\mathbf{A} = \begin{bmatrix} -1 & 3 & 1 \\ 2 & 0 & -2 \\ 0 & 4 & 5 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 0 & 2 & 3 \\ 8 & -1 & 9 \\ -2 & 0 & 5 \end{bmatrix}$$

then

$$\mathbf{AB} = \begin{bmatrix} -1 & 3 & 1 \\ 2 & 0 & -2 \\ 0 & 4 & 5 \end{bmatrix}_{3 \times 3} \begin{bmatrix} 0 & 2 & 3 \\ 8 & -1 & 9 \\ -2 & 0 & 5 \end{bmatrix}_{3 \times 3} = \begin{bmatrix} 22 & -5 & 29 \\ 4 & 4 & -4 \\ 22 & -4 & 61 \end{bmatrix}_{3 \times 3}$$

$$\mathbf{BA} = \begin{bmatrix} 0 & 2 & 3 \\ 8 & -1 & 9 \\ -2 & 0 & 5 \end{bmatrix}_{3 \times 3} \begin{bmatrix} -1 & 3 & 1 \\ 2 & 0 & -2 \\ 0 & 4 & 5 \end{bmatrix}_{3 \times 3} = \begin{bmatrix} 4 & 12 & 11 \\ -10 & 60 & 55 \\ 2 & 14 & 23 \end{bmatrix}_{3 \times 3}$$

NOTE: When a row vector premultiplies a column vector, the result is their inner product, a scalar whose value is the sum of the products of the elements of the two vectors. When an $n \times 1$ column vector premultiplies a $1 \times n$ row vector, the result is an $n \times n$ square matrix whose elements are the products of the elements of the vectors; thus if

$$\mathbf{u} = [u_1, \dots, u_n] \quad \text{and} \quad \mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

then, as above, $\mathbf{u}_{1 \times n} \mathbf{v}_{n \times 1} = w$ (a scalar) where $w = \sum_{i=1}^n u_i v_i$, and $\mathbf{v}_{n \times 1} \mathbf{u}_{1 \times n} = \mathbf{x}_{n \times n}$ (a square matrix), where $x_{ij} = v_i u_j$.

EXAMPLE

$$(a) [1, 3, 6] \begin{bmatrix} -2 \\ 4 \\ -1 \end{bmatrix} = -2 + 12 - 6 = 4$$

$$\begin{bmatrix} -2 \\ 4 \\ -1 \end{bmatrix} [1, 3, 6] = \begin{bmatrix} -2 & -6 & 12 \\ 4 & 12 & 24 \\ -1 & -3 & -6 \end{bmatrix}$$

$$(b) [1, 0, 0, 2, -3] \begin{bmatrix} 0 \\ -1 \\ -4 \\ -3 \\ -2 \end{bmatrix} = 0 + 0 + 0 - 6 + 6 = 0$$

$$\begin{bmatrix} 0 \\ -1 \\ -4 \\ -3 \\ -2 \end{bmatrix} [1, 0, 0, 2, -3] = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & -2 & 3 \\ -4 & 0 & 0 & -8 & 12 \\ -3 & 0 & 0 & -6 & 9 \\ -2 & 0 & 0 & -4 & 6 \end{bmatrix}$$

Although the sequence in which two matrices are multiplied affects the result, the order in which three or more matrices are multiplied does not affect the result, provided the sequence is preserved. That is,

$$\mathbf{A}_{m \times n} \mathbf{B}_{n \times p} \mathbf{C}_{p \times q} = \mathbf{A}_{m \times n} (\mathbf{B}_{n \times p} \mathbf{C}_{p \times q}) = (\mathbf{A}_{m \times n} \mathbf{B}_{n \times p}) \mathbf{C}_{p \times q}$$

A corresponding property holds for multiplication of numbers. In summary, addition of matrices is commutative, that is, $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$ and both addition and subtraction are associative, that is, $\mathbf{A} \pm \mathbf{B} \pm \mathbf{C} = \mathbf{A} \pm (\mathbf{B} \pm \mathbf{C}) = (\mathbf{A} \pm \mathbf{B}) \pm \mathbf{C}$; multiplication of matrices is *not* commutative, that is, $\mathbf{AB} \neq \mathbf{BA}$, but is associative, that is, $\mathbf{ABC} = \mathbf{A(BC)} = (\mathbf{AB})\mathbf{C}$. For numbers, addition, subtraction, and multiplication are both associative and commutative.

EXAMPLE

$$\begin{array}{ccccccc} \mathbf{A} & & \mathbf{B} & & \mathbf{C} & = & \mathbf{A} & & \mathbf{BC} \\ \begin{bmatrix} 1 & -1 & 0 \\ 2 & -3 & 4 \\ 0 & 0 & 1 \end{bmatrix}_{3 \times 3} & & \begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix}_{3 \times 1} & & [0, 2, -1]_{1 \times 3} & = & \begin{bmatrix} 1 & -1 & 0 \\ 2 & -3 & 4 \\ 0 & 0 & 1 \end{bmatrix}_{3 \times 3} & & \begin{bmatrix} 0 & 6 & -3 \\ 0 & 2 & -1 \\ 0 & -2 & 1 \end{bmatrix}_{3 \times 3} \\ & & & & & & = & \mathbf{ABC} & \\ & & & & & & = & \begin{bmatrix} 0 & 4 & -2 \\ 0 & -2 & 1 \\ 0 & -2 & 1 \end{bmatrix}_{3 \times 3} \end{array}$$

Alternatively, by associativity,

$$\begin{array}{ccccccc} \mathbf{A} & & \mathbf{B} & & \mathbf{C} & = & \mathbf{AB} & & \mathbf{C} \\ \begin{bmatrix} 1 & -1 & 0 \\ 2 & -3 & 4 \\ 0 & 0 & 1 \end{bmatrix}_{3 \times 3} & & \begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix}_{3 \times 1} & & [0, 2, -1]_{1 \times 3} & = & \begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix}_{3 \times 1} & & [0, 2, -1]_{1 \times 3} \\ & & & & & & = & \mathbf{ABC} & \\ & & & & & & = & \begin{bmatrix} 0 & 4 & -2 \\ 0 & -2 & 1 \\ 0 & -2 & 1 \end{bmatrix}_{3 \times 3} \end{array}$$

EXAMPLE

$$\begin{array}{ccccccc} \mathbf{A} & & \mathbf{B} & & \mathbf{C} & = & \mathbf{A} & & \mathbf{BC} \\ \begin{bmatrix} 1 & -1 \\ 6 & 10 \end{bmatrix}_{2 \times 2} & & \begin{bmatrix} 5 & -2 & 3 \\ -8 & 0 & 6 \end{bmatrix}_{2 \times 3} & & \begin{bmatrix} -1 \\ 0 \\ -4 \end{bmatrix}_{3 \times 1} & = & \begin{bmatrix} 1 & -1 \\ 6 & 10 \end{bmatrix}_{2 \times 2} & & \begin{bmatrix} -17 \\ -16 \end{bmatrix}_{2 \times 1} \\ & & & & & & = & \mathbf{ABC} & \\ & & & & & & = & \begin{bmatrix} -1 \\ -262 \end{bmatrix}_{2 \times 1} \end{array}$$

Alternatively, by associativity,

$$\begin{array}{ccccccc}
 \mathbf{A} & & \mathbf{B} & & \mathbf{C} & = & \mathbf{A} \mathbf{B} & & \mathbf{C} \\
 \left[\begin{array}{cc} 1 & -1 \\ 6 & 10 \end{array} \right]_{2 \times 2} & & \left[\begin{array}{cc} 5 & -2 \\ -8 & 0 \end{array} \right]_{2 \times 3} & & \left[\begin{array}{c} -1 \\ 0 \\ -4 \end{array} \right]_{3 \times 1} & = & \left[\begin{array}{ccc} 13 & -2 & -3 \\ -50 & -12 & 78 \end{array} \right]_{2 \times 3} & & \left[\begin{array}{c} -1 \\ 0 \\ -4 \end{array} \right]_{3 \times 1} \\
 & & & & & = & \mathbf{A} \mathbf{B} \mathbf{C} \\
 & & & & & = & \left[\begin{array}{c} -1 \\ -262 \end{array} \right]_{2 \times 1}
 \end{array}$$

EXAMPLE

$$\begin{array}{ccccccc}
 \mathbf{A} & & \mathbf{B} & & \mathbf{C} & = & \mathbf{A} & & \mathbf{B} \mathbf{C} \\
 \left[\begin{array}{cc} 3 & 4 \\ 4 & 3 \end{array} \right]_{2 \times 2} & & \left[\begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array} \right]_{2 \times 2} & & \left[\begin{array}{ccc} -6 & -8 & 10 \\ -5 & 5 & 4 \end{array} \right]_{2 \times 3} & = & \left[\begin{array}{cc} 3 & 4 \\ 4 & 3 \end{array} \right]_{2 \times 2} & & \left[\begin{array}{ccc} -5 & 5 & 4 \\ 6 & 8 & -10 \end{array} \right]_{2 \times 3} \\
 & & & & & = & \mathbf{A} \mathbf{B} \mathbf{C} \\
 & & & & & = & \left[\begin{array}{ccc} 9 & 47 & -28 \\ -2 & 44 & -14 \end{array} \right]_{2 \times 3}
 \end{array}$$

Alternatively, by associativity,

$$\begin{array}{ccccccc}
 \mathbf{A} & & \mathbf{B} & & \mathbf{C} & = & \mathbf{A} \mathbf{B} & & \mathbf{C} \\
 \left[\begin{array}{cc} 3 & 4 \\ 4 & 3 \end{array} \right]_{2 \times 2} & & \left[\begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array} \right]_{2 \times 2} & & \left[\begin{array}{ccc} -6 & -8 & 10 \\ -5 & 5 & 4 \end{array} \right]_{2 \times 3} & = & \left[\begin{array}{cc} -4 & 3 \\ -3 & 4 \end{array} \right]_{2 \times 2} & & \left[\begin{array}{ccc} -6 & -8 & 10 \\ -5 & 5 & 4 \end{array} \right]_{2 \times 3} \\
 & & & & & = & \mathbf{A} \mathbf{B} \mathbf{C} \\
 & & & & & = & \left[\begin{array}{ccc} 9 & 47 & -28 \\ -2 & 44 & -14 \end{array} \right]_{2 \times 3}
 \end{array}$$

EXAMPLE

$$\begin{array}{ccccccc}
 \mathbf{A} & & \mathbf{B} & & \mathbf{C} & = & \mathbf{A} & & \mathbf{B} \mathbf{C} \\
 [1, -1, 4, 2]_{1 \times 4} & & \left[\begin{array}{c} 0 \\ 2 \\ -1 \\ 0 \end{array} \right]_{4 \times 1} & & [-1, 0, 6]_{1 \times 3} & = & [1, -1, 4, 2]_{1 \times 4} & & \left[\begin{array}{ccc} 0 & 0 & 0 \\ -2 & 0 & 12 \\ 1 & 0 & -6 \\ 0 & 0 & 0 \end{array} \right]_{4 \times 3} \\
 & & & & & = & \mathbf{A} \mathbf{B} \mathbf{C} \\
 & & & & & = & [6, 0, -36]_{1 \times 3}
 \end{array}$$

Alternatively, by associativity,

$$\begin{array}{ccccccc}
 \mathbf{A} & & \mathbf{B} & & \mathbf{C} & = & \mathbf{A} \mathbf{B} & & \mathbf{C} \\
 [1, -1, 4, 2]_{1 \times 4} & & \left[\begin{array}{c} 0 \\ 2 \\ -1 \\ 0 \end{array} \right]_{4 \times 1} & & [-1, 0, 6]_{1 \times 3} & = & -6 \times [-1, 0, 6]_{1 \times 3} \\
 & & & & & = & \mathbf{A} \mathbf{B} \mathbf{C} \\
 & & & & & = & [6, 0, -36]_{1 \times 3}
 \end{array}$$

PROBLEMS

Obtain the matrix resulting from each of the following operations.

$$1. 3 \begin{bmatrix} -1 & 2 & 3 & 4 \\ 0 & -1 & -1 & 2 \\ 1 & 2 & 0 & 3 \end{bmatrix}$$

$$2. \begin{bmatrix} 2 & -1 \\ 1 & 0 \\ 3 & -3 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}$$

$$3. [1, -1, 0, 2] \begin{bmatrix} 2 & 1 \\ 1 & 3 \\ -1 & 0 \\ 0 & 2 \end{bmatrix}$$

$$4. [1, -2, 0] \begin{bmatrix} 1 & -1 \\ 3 & 0 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$5. \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$6. [1, -1, 1] \begin{bmatrix} 2 & 0 \\ 0 & 2 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Answers to Odd-Numbered Problems

$$1. \begin{bmatrix} -3 & 6 & 9 & 12 \\ 0 & -3 & -3 & 6 \\ 3 & 6 & 0 & 9 \end{bmatrix}$$

$$3. [1, 2]$$

$$5. \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

7.4 Special Types of Matrices

In this section three special types of matrices are discussed. These are diagonal matrices, identity matrices, and null matrices.

DIAGONAL MATRICES

A *diagonal matrix* is a square matrix that has zeros everywhere except on the main diagonal, that is, the diagonal running from upper left to lower right; thus

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} = (a_{ij})_{n \times n}$$

is a diagonal matrix if and only if

$$a_{ij} = 0 \quad \text{for } i \neq j$$

$$a_{ij} \neq 0 \quad \text{for at least one } i = j \text{ (if all the elements of a matrix are zero, it is a null matrix discussed below)}$$

An $n \times n$ diagonal matrix may be indicated by the notation

$$\mathbf{A} = \begin{bmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{bmatrix}$$

or by

$$\mathbf{D}_n = \begin{bmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{bmatrix}$$

EXAMPLE

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 10 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 7 \end{bmatrix}$$

are diagonal matrices.

IDENTITY MATRICES

An *identity matrix* is a diagonal matrix each of whose diagonal elements is positive one; thus

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} = (a_{ij})_{n \times n}$$

is an identity matrix if and only if

$$a_{ij} = 0 \quad \text{for } i \neq j$$

$$a_{ij} = 1 \quad \text{for } i = j$$

or equivalently, a diagonal matrix \mathbf{D}_n is an identity matrix if and only if

$$d_i = 1 \quad \text{for } i = 1, 2, \dots, n$$

An $n \times n$ identity matrix is denoted by \mathbf{I}_n .

EXAMPLE

$$\mathbf{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{I}_6 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

are the 3×3 and 6×6 identity matrices.

Note that premultiplying or postmultiplying a matrix by the appropriately sized identity matrix leaves it unchanged—that is,

$$\mathbf{A}_{m \times n} = \mathbf{I}_m \mathbf{A}_{m \times n} = \mathbf{A}_{m \times n} \mathbf{I}_n = \mathbf{A}_{m \times n}$$

EXAMPLE

$$\begin{aligned}
& \begin{matrix} A_{2 \times 4} \\ \left[\begin{array}{cccc} 2 & -3 & -6 & 4 \\ -1 & -1 & 0 & 3 \end{array} \right]_{2 \times 4} \end{matrix} &= \begin{matrix} I_2 \\ \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right]_{2 \times 2} \end{matrix} \begin{matrix} A_{2 \times 4} \\ \left[\begin{array}{cccc} 2 & -3 & -6 & 4 \\ -1 & -1 & 0 & 3 \end{array} \right]_{2 \times 4} \end{matrix} \\
& &= \begin{matrix} A_{2 \times 4} \\ \left[\begin{array}{cccc} 2 & -3 & -6 & 4 \\ -1 & -1 & 0 & 3 \end{array} \right]_{2 \times 4} \end{matrix} \begin{matrix} I_4 \\ \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right]_{4 \times 4} \end{matrix} \\
& &= \begin{matrix} A_{2 \times 4} \\ \left[\begin{array}{cccc} 2 & -3 & -6 & 4 \\ -1 & -1 & 0 & 3 \end{array} \right]_{2 \times 4} \end{matrix}
\end{aligned}$$

EXAMPLE

$$\begin{aligned}
& \begin{matrix} A_{3 \times 3} \\ \left[\begin{array}{ccc} 3 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 6 \end{array} \right]_{3 \times 3} \end{matrix} &= \begin{matrix} I_3 \\ \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right]_{3 \times 3} \end{matrix} \begin{matrix} A_{3 \times 3} \\ \left[\begin{array}{ccc} 3 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 6 \end{array} \right]_{3 \times 3} \end{matrix} \\
& &= \begin{matrix} A_{3 \times 3} \\ \left[\begin{array}{ccc} 3 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 6 \end{array} \right]_{3 \times 3} \end{matrix} \begin{matrix} I_3 \\ \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right]_{3 \times 3} \end{matrix} \\
& &= \begin{matrix} A_{3 \times 3} \\ \left[\begin{array}{ccc} 3 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 6 \end{array} \right]_{3 \times 3} \end{matrix}
\end{aligned}$$

NULL MATRICES

A *null matrix* is an $m \times n$ matrix all of whose elements are zeros; it is denoted by \mathbf{O} . When the appropriately sized null matrix is added to or subtracted from another matrix that matrix is unchanged, that is,

$$\mathbf{A}_{m \times n} \pm \mathbf{O}_{m \times n} = \mathbf{A}_{m \times n}$$

Premultiplying or postmultiplying a matrix by the appropriately sized null matrix results in another null matrix—that is,

$$\mathbf{O}_{k \times m} \mathbf{A}_{m \times n} = \mathbf{O}_{k \times n} \quad \text{and} \quad \mathbf{A}_{m \times n} \mathbf{O}_{n \times 1} = \mathbf{O}_{m \times 1}$$

EXAMPLE

$$\begin{aligned}
& \begin{matrix} A_{2 \times 3} \\ \left[\begin{array}{ccc} 0 & 1 & 0 \\ 10 & 0 & -3 \end{array} \right]_{2 \times 3} \end{matrix} \pm \begin{matrix} O_{2 \times 3} \\ \left[\begin{array}{ccc} 0 & 0 & 0 \\ 0 & 0 & 0 \end{array} \right]_{2 \times 3} \end{matrix} = \begin{matrix} A_{2 \times 3} \\ \left[\begin{array}{ccc} 0 & 1 & 0 \\ 10 & 0 & -3 \end{array} \right]_{2 \times 3} \end{matrix}
\end{aligned}$$

EXAMPLE

$$\begin{array}{ccc}
\mathbf{O}_{3 \times 2} & \mathbf{A}_{2 \times 4} & = \mathbf{O}_{3 \times 4} \\
\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}_{3 \times 2} & \begin{bmatrix} 1 & 6 & 4 & -10 \\ -1 & 20 & -13 & 11 \end{bmatrix}_{2 \times 4} & = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}_{3 \times 4} \\
\mathbf{A}_{2 \times 4} & \mathbf{O}_{4 \times 1} & = \mathbf{O}_{2 \times 1} \\
\begin{bmatrix} 1 & 6 & 4 & -10 \\ -1 & 20 & -13 & 11 \end{bmatrix}_{2 \times 4} & \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}_{4 \times 1} & = \begin{bmatrix} 0 \\ 0 \end{bmatrix}_{2 \times 1}
\end{array}$$

7.5 The Transpose of a Matrix

In many applications, the transpose of a matrix is used. The transpose of matrix is defined in this section and the transpose of a sum or difference of matrices and the transpose of a product of matrices are discussed.

The *transpose* of an $m \times n$ matrix \mathbf{A} is an $n \times m$ matrix, denoted by \mathbf{A}' , whose rows are the columns of \mathbf{A} and whose columns are the rows of \mathbf{A} . Thus if

$$\mathbf{A}_{m \times n} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = (a_{ij})_{m \times n}$$

then the transpose of \mathbf{A} is

$$\mathbf{A}'_{n \times m} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}' = \begin{bmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & & \vdots \\ a_{1n} & \cdots & a_{mn} \end{bmatrix} = (a_{ji})_{n \times m} = (a_{ij})'_{m \times n}$$

Note that the transpose of an n -dimensional row vector is an n -dimensional column vector and, similarly, the transpose of an n -dimensional column vector is an n -dimensional row vector. The transpose of a diagonal matrix is that same diagonal matrix.

EXAMPLES

$$\text{If } \mathbf{A} = \begin{bmatrix} 3 & 2 & 0 \\ -6 & 0 & -1 \end{bmatrix}_{2 \times 3}, \text{ then } \mathbf{A}' = \begin{bmatrix} 3 & -6 \\ 2 & 0 \\ 0 & -1 \end{bmatrix}_{3 \times 2}$$

$$\text{If } \mathbf{A} = [6, 2, -1, 0, -5]_{1 \times 5}, \text{ then } \mathbf{A}' = \begin{bmatrix} 6 \\ 2 \\ -1 \\ 0 \\ -5 \end{bmatrix}_{5 \times 1}$$

$$\text{If } \mathbf{A} = \begin{bmatrix} 6 & -8 & -10 \\ -2 & 0 & 7 \\ 30 & 40 & 0 \end{bmatrix}_{3 \times 3}, \text{ then } \mathbf{A}' = \begin{bmatrix} 6 & -2 & 30 \\ -8 & 0 & 40 \\ -10 & 7 & 0 \end{bmatrix}_{3 \times 3}$$

$$\text{If } A = \begin{bmatrix} 3 \\ 0 \\ -1 \\ 4 \end{bmatrix}_{4 \times 1}, \text{ then } A' = [3, 0, -1, 4]_{1 \times 4}$$

$$\text{If } A = \begin{bmatrix} -1 & 5 & -3 \\ 5 & 0 & 4 \\ -3 & 4 & 9 \end{bmatrix}_{3 \times 3}, \text{ then } A' = \begin{bmatrix} -1 & 5 & -3 \\ 5 & 0 & 4 \\ -3 & 4 & 9 \end{bmatrix}_{3 \times 3}$$

$$\text{If } A = \begin{bmatrix} 6 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 5 \end{bmatrix}_{4 \times 4}, \text{ then } A' = \begin{bmatrix} -6 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 5 \end{bmatrix}_{4 \times 4}$$

If a (square) matrix and its transpose are equal, that is, if $a_{ij} = a_{ji}$ for all i and j , then the matrix is said to be *symmetric* (about its main diagonal). In the last two examples above, the matrices are symmetric.

A symmetric matrix that reproduces itself when it is multiplied by itself is said to be *idempotent*. Thus A is idempotent if and only if

$$A' = A$$

and

$$A A = A$$

EXAMPLE

The identity matrix (of any size) is idempotent, since

$$I_n' = I_n$$

and

$$I_n I_n = I_n$$

EXAMPLE

The matrix

$$\begin{bmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{4}{5} \end{bmatrix}$$

is idempotent, since

$$\begin{bmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{4}{5} \end{bmatrix}' = \begin{bmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{4}{5} \end{bmatrix}$$

and

$$\begin{bmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{4}{5} \end{bmatrix} \begin{bmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{4}{5} \end{bmatrix} = \begin{bmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{4}{5} \end{bmatrix}$$

TRANSPOSE OF A SUM OR DIFFERENCE OF MATRICES

The transpose of a sum or difference of matrices is equal to the sum or difference of the transposes of the matrices; thus

$$(\mathbf{A}_{m \times n} \pm \mathbf{B}_{m \times n} \pm \mathbf{C}_{m \times n})' = \mathbf{A}'_{n \times m} \pm \mathbf{B}'_{n \times m} \pm \mathbf{C}'_{n \times m}$$

that is,

$$(d_{ij})'_{m \times n} = (d_{ji})_{n \times m}$$

where $d_{ij} = a_{ij} \pm b_{ij} \pm c_{ij}$ and $d_{ji} = a_{ji} \pm b_{ji} \pm c_{ji}$.

EXAMPLE

If

$$\mathbf{A} = \begin{bmatrix} 3 & 0 & 2 \\ -6 & 8 & -1 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} -1 & 6 & -2 \\ 3 & -2 & 1 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 0 & 0 & 6 \\ -2 & 5 & 25 \end{bmatrix}$$

then

$$[\mathbf{A} + \mathbf{B} + \mathbf{C}]' = \begin{bmatrix} 2 & 6 & 6 \\ -5 & 11 & 25 \end{bmatrix}' = \begin{bmatrix} 2 & -5 \\ 6 & 11 \\ 6 & 25 \end{bmatrix}$$

or, alternatively,

$$\mathbf{A}' + \mathbf{B}' + \mathbf{C}' = \begin{bmatrix} 3 & -6 \\ 0 & 8 \\ 2 & -1 \end{bmatrix} + \begin{bmatrix} -1 & 3 \\ 6 & -2 \\ -2 & 1 \end{bmatrix} + \begin{bmatrix} 0 & -2 \\ 0 & 5 \\ 6 & 25 \end{bmatrix} = \begin{bmatrix} 2 & -5 \\ 6 & 11 \\ 6 & 25 \end{bmatrix}$$

EXAMPLE

If

$$\mathbf{A} = \begin{bmatrix} 2 & -1 \\ 0 & -1 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 6 & 8 \\ 10 & -12 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix}$$

then

$$[\mathbf{A} - \mathbf{B} + \mathbf{C}]' = \begin{bmatrix} -4 & -10 \\ -10 & 11 \end{bmatrix}' = \begin{bmatrix} -4 & -10 \\ -10 & 11 \end{bmatrix}$$

or, alternatively,

$$\mathbf{A}' - \mathbf{B}' + \mathbf{C}' = \begin{bmatrix} 2 & 0 \\ -1 & -1 \end{bmatrix} - \begin{bmatrix} 6 & 10 \\ 8 & -12 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} -4 & -10 \\ -10 & 11 \end{bmatrix}$$

TRANSPOSE OF A PRODUCT OF MATRICES

The transpose of a product of matrices is equal to the product of the transposes of the matrices in reverse sequence; thus

$$[\mathbf{A}_{m \times n} \mathbf{B}_{n \times p} \mathbf{C}_{p \times q}]' = \mathbf{C}'_{q \times p} \mathbf{B}'_{p \times n} \mathbf{A}'_{n \times m}$$

EXAMPLE

If

$$\mathbf{A} = \begin{bmatrix} 3 & 0 \\ -4 & -1 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 3 & 5 & -7 \\ 0 & -1 & 8 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 6 \\ -1 \\ 0 \end{bmatrix}$$

then

$$\mathbf{ABC} = \begin{bmatrix} 39 \\ -53 \end{bmatrix}$$

$$\begin{aligned} [\mathbf{ABC}]' &= \begin{bmatrix} 39 \\ -53 \end{bmatrix}' \\ &= [39, -53] \end{aligned}$$

or, alternatively,

$$\begin{aligned} [\mathbf{ABC}]' &= \mathbf{C}'\mathbf{B}'\mathbf{A}' = [6, -1, 0] \begin{bmatrix} 3 & 0 \\ 5 & -1 \\ -7 & 8 \end{bmatrix} \begin{bmatrix} 3 & -4 \\ 0 & -1 \end{bmatrix} \\ &= [39, -53] \end{aligned}$$

EXAMPLE

If

$$\mathbf{A} = [3, -1, 0] \quad \mathbf{B} = \begin{bmatrix} 6 & 0 \\ -7 & 2 \\ 0 & 3 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

then $\mathbf{ABC} = -2$ and $[\mathbf{ABC}]' = -2$ (the transpose of a scalar is that scalar) or, alternatively,

$$[\mathbf{ABC}]' = \mathbf{C}'\mathbf{B}'\mathbf{A}' = [0, 1] \begin{bmatrix} 6 & -7 & 0 \\ 0 & 2 & 3 \end{bmatrix} \begin{bmatrix} 3 \\ -1 \\ 0 \end{bmatrix} = -2$$

PROBLEMS

Obtain the transpose of each of the following matrices.

1. $\begin{bmatrix} 2 & 6 \\ 5 & -3 \end{bmatrix}$

2. $\begin{bmatrix} -3 \\ 6 \\ 5 \\ -2 \end{bmatrix}$

3. $\begin{bmatrix} -7 & -5 & 6 \\ 8 & 10 & -3 \end{bmatrix}$

4. $\begin{bmatrix} 1 & -6 & 3 \\ -1 & 4 & -8 \\ 0 & 5 & 7 \end{bmatrix}$

5. $[3, -2, 4, -9]$

6. $\begin{bmatrix} -1 & 3 \\ 2 & 1 \\ -4 & 2 \end{bmatrix}$

7. If $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$, $B = \begin{bmatrix} -2 & 3 \\ 4 & 0 \end{bmatrix}$, $C = \begin{bmatrix} 0 & -1 \\ 2 & 3 \end{bmatrix}$; find:

- $[A + B]'$
- $[B + C]'$
- $[AB]'$
- $[ABC]'$

8. If $A = \begin{bmatrix} 2 & -1 \\ 1 & 2 \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$, $C = \begin{bmatrix} 3 & 0 & 2 \\ -1 & 2 & -4 \end{bmatrix}$,

and $D = \begin{bmatrix} -3 & 0 & -2 \\ 1 & -2 & 4 \end{bmatrix}$, find:

- $[AB]'$
- $[C + D]'$
- $[D'C]'$
- $A + D' + C'$

Answers to Odd-Numbered Problems

1. $\begin{bmatrix} 2 & 5 \\ 6 & -3 \end{bmatrix}$

3. $\begin{bmatrix} -7 & 8 \\ -5 & 10 \\ 6 & -3 \end{bmatrix}$

5. $\begin{bmatrix} 3 \\ -2 \\ 4 \\ -9 \end{bmatrix}$

7. a. $\begin{bmatrix} -1 & 6 \\ 5 & 1 \end{bmatrix}$

b. $\begin{bmatrix} -2 & 6 \\ 2 & 3 \end{bmatrix}$

c. $\begin{bmatrix} 6 & 0 \\ 3 & 6 \end{bmatrix}$

d. $\begin{bmatrix} 6 & 12 \\ 3 & 18 \end{bmatrix}$

7.6 Partitioned Matrices

It is frequently convenient to partition a matrix into submatrices. These submatrices can then be treated as scalars in performing operations on the original matrix. Partitioning of a matrix is indicated by horizontal and vertical lines between rows and columns.

The $m \times n$ matrix A may be partitioned, for example, as follows

$$A = (A_1 \mid A_2)$$

where A_1 is $m \times n_1$, A_2 is $m \times n_2$, and $n_1 + n_2 = n$. The transpose of a partitioned matrix can be written in terms of the transposes of its submatrices; thus

$$A' = \begin{bmatrix} A_1' \\ A_2' \end{bmatrix}$$

EXAMPLE

If

$$\mathbf{A} = [\mathbf{A}_1 \mid \mathbf{A}_2] = \left[\begin{array}{cc|cc} 4 & -3 & 5 & 0 \\ 2 & -1 & 1 & 6 \\ 8 & -2 & 3 & -7 \end{array} \right]$$

then

$$\mathbf{A}' = \left[\begin{array}{c} \mathbf{A}'_1 \\ \mathbf{A}'_2 \end{array} \right] = \left[\begin{array}{ccc|ccc} 4 & 2 & 8 & & & \\ -3 & -1 & -2 & & & \\ \hline 5 & 1 & 3 & & & \\ 0 & 6 & -7 & & & \end{array} \right]$$

If they are partitioned conformably, partitioned matrices can be added, subtracted, or multiplied. If an $m \times n$ matrix \mathbf{A} is partitioned $\mathbf{A} = [\mathbf{A}_1 \mid \mathbf{A}_2]$, where \mathbf{A}_1 is $m \times n_1$, \mathbf{A}_2 is $m \times n_2$, and $n_1 + n_2 = n$, and an $m \times n$ matrix \mathbf{B} is partitioned $\mathbf{B} = [\mathbf{B}_1 \mid \mathbf{B}_2]$, where \mathbf{B}_1 is $m \times n_1$, \mathbf{B}_2 is $m \times n_2$, and $n_1 + n_2 = n$, then

$$\mathbf{A} \pm \mathbf{B} = [\mathbf{A}_1 \pm \mathbf{B}_1 \mid \mathbf{A}_2 \pm \mathbf{B}_2]$$

Similarly, if

$$\mathbf{A} = \left[\begin{array}{c} \mathbf{A}_1 \\ \mathbf{A}_2 \end{array} \right]$$

where \mathbf{A} is $m \times n$, \mathbf{A}_1 is $m_1 \times n$, \mathbf{A}_2 is $m_2 \times n$, and $m_1 + m_2 = m$, and

$$\mathbf{B} = \left[\begin{array}{c} \mathbf{B}_1 \\ \mathbf{B}_2 \end{array} \right]$$

where \mathbf{B} is $m \times n$, \mathbf{B}_1 is $m_1 \times n$, \mathbf{B}_2 is $m_2 \times n$, and $m_1 + m_2 = m$, then

$$\mathbf{A} \pm \mathbf{B} = \left[\begin{array}{c} \mathbf{A}_1 \pm \mathbf{B}_1 \\ \mathbf{A}_2 \pm \mathbf{B}_2 \end{array} \right]$$

EXAMPLE

If

$$\mathbf{A} = \left[\begin{array}{cccc} 3 & 4 & -2 & 0 \\ -1 & 0 & 5 & 6 \\ 7 & 3 & 3 & 2 \end{array} \right] \quad \text{and} \quad \mathbf{B} = \left[\begin{array}{cccc} -3 & 0 & 2 & 4 \\ -1 & -1 & -2 & 5 \\ 5 & 4 & 3 & 1 \end{array} \right]$$

then

$$\begin{aligned} \mathbf{A} + \mathbf{B} &= (\mathbf{A}_1 \mid \mathbf{A}_2) + (\mathbf{B}_1 \mid \mathbf{B}_2) \\ &= \left[\begin{array}{cc|cc} 3 & 4 & -2 & 0 \\ -1 & 0 & 5 & 6 \\ 7 & 3 & 3 & 2 \end{array} \right] + \left[\begin{array}{cc|cc} -3 & 0 & 2 & 4 \\ -1 & -1 & -2 & 5 \\ 5 & 4 & 3 & 1 \end{array} \right] \\ &= \left[\begin{array}{cc|cc} 0 & 4 & 0 & 4 \\ -2 & -1 & 3 & 11 \\ 12 & 7 & 6 & 3 \end{array} \right] = [\mathbf{A}_1 + \mathbf{B}_1 \mid \mathbf{A}_2 + \mathbf{B}_2] \end{aligned}$$

or

$$\begin{aligned}
 \mathbf{A} + \mathbf{B} &= \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{bmatrix} \\
 &= \begin{bmatrix} 3 & 4 & -2 & 0 \\ -1 & 0 & 5 & 6 \\ 7 & 3 & 3 & 2 \end{bmatrix} + \begin{bmatrix} -3 & 0 & 2 & 4 \\ -1 & -1 & -2 & 5 \\ 5 & 4 & 3 & 1 \end{bmatrix} \\
 &= \begin{bmatrix} 0 & 4 & 0 & 4 \\ -2 & -1 & 3 & 11 \\ 12 & 7 & 6 & 3 \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 + \mathbf{B}_1 \\ \mathbf{A}_2 + \mathbf{B}_2 \end{bmatrix}
 \end{aligned}$$

Many other partitionings are possible and will give the same result.

Partitioning is frequently conceptually convenient when matrices are to be added or subtracted, but the computational advantages of matrix partitioning are associated primarily with multiplication and other more complex operations. Partitioned matrices must be partitioned conformably for multiplication. If an $m \times n$ matrix \mathbf{A} is partitioned $\mathbf{A} = [\mathbf{A}_1 \mid \mathbf{A}_2]$, where \mathbf{A}_1 is $m \times n_1$, \mathbf{A}_2 is $m \times n_2$, and $n_1 + n_2 = n$, and an $n \times p$ matrix \mathbf{B} is partitioned

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{bmatrix}$$

where \mathbf{B}_1 is $n_1 \times p$ and \mathbf{B}_2 is $n_2 \times p$, then

$$\mathbf{AB} = [\mathbf{A}_1 \mid \mathbf{A}_2] \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{bmatrix} = \mathbf{A}_1\mathbf{B}_1 + \mathbf{A}_2\mathbf{B}_2$$

EXAMPLE

If

$$\mathbf{A} = \begin{bmatrix} 3 & 0 & -1 \\ -2 & 4 & 1 \\ 1 & -1 & 2 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 2 & 1 \\ 1 & 3 \\ -1 & 1 \end{bmatrix}$$

then

$$\begin{aligned}
 \mathbf{AB} &= \begin{bmatrix} 3 & 0 & -1 \\ -2 & 4 & 1 \\ 1 & -1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 3 \\ -1 & 1 \end{bmatrix} \\
 &= \begin{bmatrix} 3 & 0 \\ -2 & 4 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} + \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} \begin{bmatrix} -1 & 1 \end{bmatrix} \\
 &= \begin{bmatrix} 6 & 3 \\ 0 & 10 \\ 1 & -2 \end{bmatrix} + \begin{bmatrix} 1 & -1 \\ -1 & 1 \\ -2 & 2 \end{bmatrix} \\
 &= \begin{bmatrix} 7 & 2 \\ -1 & 11 \\ -1 & 0 \end{bmatrix}
 \end{aligned}$$

which can be verified by direct matrix multiplication.

Matrices may be partitioned into more than two submatrices. In fact, it is possible to partition an $m \times n$ matrix into a maximum of mn submatrices; note that maximum partitioning is equivalent to no partitioning, since each element is treated as a scalar matrix. Unless there is a logical partitioning in terms of the variables in a problem, matrices should be partitioned for ease of computation; this involves some reasonable compromise between minimizing the number of submatrices and minimizing their maximum size.

Frequently matrices are partitioned once horizontally and once vertically. Thus the $m \times n$ matrix \mathbf{A} may be partitioned

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}$$

where \mathbf{A}_{11} is $m_1 \times n_1$, \mathbf{A}_{12} is $m_1 \times n_2$, \mathbf{A}_{21} is $m_2 \times n_1$, \mathbf{A}_{22} is $m_2 \times n_2$, and $m_1 + m_2 = m$, $n_1 + n_2 = n$. Then

$$\mathbf{A}' = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}' = \begin{bmatrix} \mathbf{A}'_{11} & \mathbf{A}'_{12} \\ \mathbf{A}'_{21} & \mathbf{A}'_{22} \end{bmatrix}$$

If the $n \times p$ matrix \mathbf{B} is partitioned

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix}$$

where \mathbf{B}_{11} is $n_1 \times p_1$, \mathbf{B}_{12} is $n_1 \times p_2$, \mathbf{B}_{21} is $n_2 \times p_1$, \mathbf{B}_{22} is $n_2 \times p_2$, and $n_1 + n_2 = n$, $p_1 + p_2 = p$, then \mathbf{A} and \mathbf{B} are partitioned conformably for multiplication and

$$\begin{aligned} \mathbf{AB} &= \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} & \mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} \\ \mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} & \mathbf{A}_{21}\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} \end{bmatrix} \end{aligned}$$

EXAMPLE

If

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 1 \\ -1 & 0 & 1 \\ 2 & -1 & 4 \\ 0 & 2 & -3 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 3 & -2 & 1 & 0 & -1 \\ 5 & -1 & 4 & -3 & 2 \\ 3 & -2 & 0 & 1 & -1 \end{bmatrix}$$

then

$$\begin{aligned} \mathbf{AB} &= \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix} \\ &= \begin{bmatrix} 1 & 3 & 1 \\ -1 & 0 & 1 \\ 2 & -1 & 4 \\ 0 & 2 & -3 \end{bmatrix} \begin{bmatrix} 3 & -2 & 1 & 0 & -1 \\ 5 & -1 & 4 & -3 & 2 \\ 3 & -2 & 0 & 1 & -1 \end{bmatrix} \end{aligned}$$

and

$$\begin{aligned} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} &= \begin{bmatrix} 1 & 3 \\ -1 & 0 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 3 & -2 \\ 5 & -1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix} [3, -2] \\ &= \begin{bmatrix} 18 & -5 \\ -3 & 2 \\ 1 & -3 \end{bmatrix} + \begin{bmatrix} 3 & -2 \\ 3 & -2 \\ 12 & -8 \end{bmatrix} \\ &= \begin{bmatrix} 21 & -7 \\ 0 & 0 \\ 13 & -11 \end{bmatrix} \end{aligned}$$

$$\begin{aligned}
\mathbf{A}_{11} \mathbf{B}_{12} + \mathbf{A}_{12} \mathbf{B}_{22} &= \begin{bmatrix} 1 & 3 \\ -1 & 0 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ 4 & -3 & 2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix} [0, 1, -1] \\
&= \begin{bmatrix} 13 & -9 & 5 \\ -1 & 0 & 1 \\ -2 & 3 & -4 \end{bmatrix} + \begin{bmatrix} 0 & 1 & -1 \\ 0 & 1 & -1 \\ 0 & 4 & -4 \end{bmatrix} \\
&= \begin{bmatrix} 13 & -8 & 4 \\ -1 & 1 & 0 \\ -2 & 7 & -8 \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
\mathbf{A}_{21} \mathbf{B}_{11} + \mathbf{A}_{22} \mathbf{B}_{21} &= [0, 2] \begin{bmatrix} 3 & -2 \\ 5 & -1 \end{bmatrix} + [-3][3, -2] \\
&= [10, -2] + [-9, 6] \\
&= [1, 4]
\end{aligned}$$

$$\begin{aligned}
\mathbf{A}_{21} \mathbf{B}_{12} + \mathbf{A}_{22} \mathbf{B}_{22} &= [0, 2] \begin{bmatrix} 1 & 0 & -1 \\ 4 & -3 & 2 \end{bmatrix} + [-3][0, 1, -1] \\
&= [8, -6, 4] + [0, -3, 3] \\
&= [8, -9, 7]
\end{aligned}$$

Thus

$$\begin{aligned}
\mathbf{AB} &= \left[\begin{array}{cc|ccc} \mathbf{A}_{11} \mathbf{B}_{11} + \mathbf{A}_{12} \mathbf{B}_{21} & & \mathbf{A}_{11} \mathbf{B}_{12} + \mathbf{A}_{12} \mathbf{B}_{22} \\ \mathbf{A}_{21} \mathbf{B}_{11} + \mathbf{A}_{22} \mathbf{B}_{21} & & \mathbf{A}_{21} \mathbf{B}_{12} + \mathbf{A}_{22} \mathbf{B}_{22} \end{array} \right] \\
&= \left[\begin{array}{cc|ccc} 21 & -7 & 13 & -8 & 4 \\ 0 & 0 & -1 & 1 & 0 \\ 13 & -11 & -2 & 7 & -8 \\ \hline 1 & 4 & 8 & -9 & 7 \end{array} \right]
\end{aligned}$$

which can be verified by direct matrix multiplication.

PROBLEMS

1. Compute the following.

a. $2 \begin{bmatrix} 6 & 1 \\ 0 & -3 \\ -1 & 2 \end{bmatrix} - 3 \begin{bmatrix} 4 & 2 \\ 0 & 1 \\ -5 & -1 \end{bmatrix}$

b. $\begin{bmatrix} 6 & 0 & -1 \\ 1 & -3 & 2 \end{bmatrix} \begin{bmatrix} 4 & 2 \\ 0 & 1 \\ -5 & -1 \end{bmatrix}$

Use partitioned matrices to check the results.

2. If $\mathbf{U} = [1, -1, 4]$, $\mathbf{X} = [0, 1, 2]$, $\mathbf{V} = \begin{bmatrix} 5 \\ 0 \\ 1 \end{bmatrix}$, and $\mathbf{Y} = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$, find

a. $\mathbf{UV} + \mathbf{XY}$

b. $5\mathbf{UV} + 10[\mathbf{X}(2\mathbf{V} - \mathbf{Y})]$

3. If \mathbf{A} is 2×3 , \mathbf{B} is 4×3 , \mathbf{C} is 3×3 , and \mathbf{D} is 3×2 , determine the shape of

a. \mathbf{AC}

b. \mathbf{DA}

c. \mathbf{AD}

d. \mathbf{BC}

e. \mathbf{DAC}

f. \mathbf{BCDA}

4. If

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & 1 \\ 2 & 3 & 0 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 3 & 1 & 0 \\ 1 & -1 & 2 \\ 0 & 2 & 1 \end{bmatrix}$$

find $2(\mathbf{AB} - \mathbf{BA})$ and use partitioned matrices to check the result.

5. If

$$\mathbf{A} = \begin{bmatrix} -1 & -2 & -2 \\ 1 & 2 & 1 \\ -1 & -1 & 0 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} -3 & -6 & 2 \\ 2 & 4 & -1 \\ 2 & 3 & 0 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} -5 & -8 & 0 \\ 3 & 5 & 0 \\ 1 & 2 & -1 \end{bmatrix}$$

show that

a. $\mathbf{A}^2 = \mathbf{B}^2 = \mathbf{C}^2 = \mathbf{I}$

c. $\mathbf{BC} = \mathbf{CB} = \mathbf{A}$

b. $\mathbf{AB} = \mathbf{BA} = \mathbf{C}$

d. $\mathbf{AC} = \mathbf{CA} = \mathbf{B}$

Use partitioned matrices to check the results.

6. If

$$\mathbf{A} = \begin{bmatrix} 5 & 4 & -2 \\ 4 & 5 & -2 \\ -2 & -2 & 2 \end{bmatrix}$$

show that $\mathbf{A}^2 - 11\mathbf{A} + 10\mathbf{I} = \mathbf{O}$. Use partitioned matrices to check the result.

7. If

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \quad \mathbf{D} = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}$$

find a. $\mathbf{AB} - 2\mathbf{CD}$

b. \mathbf{A}^2

c. $(\mathbf{BC})^2$

8. If

$$\mathbf{U} = [1, 0, 1] \quad \mathbf{V} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \quad \mathbf{X} = \begin{bmatrix} 4 & 2 & -1 \\ -1 & 3 & 0 \\ 0 & 1 & 1 \end{bmatrix} \quad \mathbf{Y} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

find a. \mathbf{UV}

b. $\mathbf{VU} + \mathbf{X}$

c. \mathbf{XY}

Use partitioned matrices to check the results.

9. If

$$\mathbf{U} = \begin{bmatrix} 2 & -2 & -4 \\ -1 & 3 & 4 \\ 1 & -2 & -3 \end{bmatrix} \quad \mathbf{V} = \begin{bmatrix} -1 & 2 & 4 \\ 1 & -2 & -4 \\ -1 & 2 & 4 \end{bmatrix}$$

show that

a. $\mathbf{U}^2 = \mathbf{U}$ and $\mathbf{V}^2 = \mathbf{V}$

b. $\mathbf{UV} = \mathbf{VU} = \mathbf{O}$

c. $\mathbf{U} + \mathbf{V} = \mathbf{I}$

Use partitioned matrices to check the results.

10. If

$$\mathbf{X} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ -1 & -2 & -3 \end{bmatrix}$$

show that $\mathbf{X}^2 = \mathbf{O}$. Use partitioned matrices to check the result.

11. If

$$A = \begin{bmatrix} 1 & -2 & 1 \\ 2 & 1 & -3 \\ -5 & 2 & 3 \end{bmatrix} \quad X = \begin{bmatrix} 2 & 5 & -1 & -7 \\ -2 & 1 & 3 & 4 \\ 3 & 2 & 1 & 2 \end{bmatrix}$$

$$Y = \begin{bmatrix} 3 & 6 & 0 & -6 \\ -1 & 2 & 4 & 5 \\ 4 & 3 & 2 & 3 \end{bmatrix}$$

show that $AX = AY$ (although $X \neq Y$) and use partitioned matrices to check the result.

12. If

$$A = \begin{bmatrix} 1 & -2 \\ 0 & 3 \\ 2 & -1 \end{bmatrix} \quad B = \begin{bmatrix} 0 & -2 & 3 \\ 0 & 1 & 3 \end{bmatrix} \quad C = \begin{bmatrix} -2 \\ 3 \\ -3 \end{bmatrix} \quad D = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$$

find $AB - CD'$.

Answers to Odd-Numbered Problems

1. a. $\begin{bmatrix} 0 & -4 \\ 0 & -9 \\ 13 & 7 \end{bmatrix}$

b. $\begin{bmatrix} 29 & 13 \\ -6 & -3 \end{bmatrix}$

3. a. 2×3

b. 3×3

c. 2×2

d. 4×3

e. 3×3

f. 4×3

7. a. $\begin{bmatrix} -24 & -32 \\ -24 & -32 \end{bmatrix}$

b. $\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$

c. $\begin{bmatrix} 0 & 0 \\ 16 & 16 \end{bmatrix}$

7.7 The Determinant of a Matrix

The determinant of a matrix is a scalar (number), obtained from the elements of a matrix by specified operations, which is characteristic of the matrix. Determinants are defined only for square matrices. Methods for obtaining determinants and properties of determinants are discussed in this section.

The determinant of a 2×2 matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

is given by

$$\det A = |A| = a_{11}a_{22} - a_{12}a_{21}$$

EXAMPLE

$$\text{If } A = \begin{bmatrix} 3 & -6 \\ 4 & 1 \end{bmatrix} \text{ then } |A| = \begin{vmatrix} 3 & -6 \\ 4 & 1 \end{vmatrix} = 3 - (-24) = 27$$

$$\text{If } A = \begin{bmatrix} -1 & 0 \\ 6 & 10 \end{bmatrix} \text{ then } |A| = \begin{vmatrix} -1 & 0 \\ 6 & 10 \end{vmatrix} = -10 - 0 = -10$$

Similarly, the determinant of a 3×3 matrix,

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

is given by

$$|A| = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{32}a_{21} - a_{13}a_{22}a_{31} - a_{23}a_{32}a_{11} - a_{33}a_{21}a_{12}$$

The terms can be obtained by the rule illustrated in Figure 7.1, where positive product terms are formed of elements connected by solid lines and negative product terms are formed of elements connected by dashed lines. Note that similar rules are *not* valid for higher-order matrices.

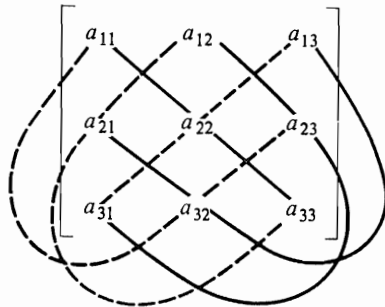
By definition, a determinant of order n , $|A| = |a_{ij}|$ with $i, j = 1, 2, \dots, n$, is the algebraic sum of $n!$ terms, each representing the product of n elements selected one from each row and one from each column in all possible and distinct combinations, having either a positive or negative sign depending on whether the number of inversions in the order of j (after the order of i in the product term has been arranged in numerical ascendancy) is even or odd.

Using this rule to obtain the determinant of a matrix larger than, say, 3×3 is tedious. For example, the determinant of a 4×4 matrix is the algebraic sum of $4! = 24$ terms, each of which is the product of four elements; the determinant of a 5×5 matrix is the algebraic sum of $5! = 120$ terms, each of which is the product of five elements.

Determinants of matrices larger than 3×3 are frequently computed by a procedure known as *expansion by cofactors*. For example, the determinant of the 3×3 matrix above can be written

$$\begin{aligned} |A| &= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \\ &= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \end{aligned}$$

Figure 7.1



Note that each determinant in the sum is the determinant of a submatrix of A obtained by deleting a particular row and column of A . These determinants are called *minors*.

DEFINITION: Let M_{ij} denote the $(n-1) \times (n-1)$ matrix obtained by deleting the i th row and the j th column of $A_{n \times n}$. The determinant

$$|M_{ij}|$$

is a *minor* of the matrix A . The scalar

$$C_{ij} = (-1)^{i+j} |M_{ij}|$$

is referred to as the *cofactor* or *signed minor* of the element a_{ij} of the matrix A . The $n \times n$ matrix

$$(C_{ij})'$$

is referred to as the *adjoint* of A and is denoted by $\text{adj } A$.

As noted above, the determinant of a matrix can be obtained by a procedure known as *expansion by cofactors*. The determinant of A can be expanded in terms of row i by the formula

$$|A| = \sum_{j=1}^n a_{ij} C_{ij} \quad \text{for any row } i = 1, 2, \dots, n$$

and in terms of column j by the formula

$$|A| = \sum_{i=1}^n a_{ij} C_{ij} \quad \text{for any column } j = 1, 2, \dots, n$$

Thus the determinant of $A_{3 \times 3}$ given above as

$$|A| = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

can be written

$$|A| = a_{11} C_{11} + a_{12} C_{12} + a_{13} C_{13} = \sum_{j=1}^3 a_{1j} C_{1j}$$

EXAMPLE

Find the determinant of the matrix

$$A = \begin{bmatrix} 3 & 0 & -2 \\ 6 & -8 & 1 \\ 0 & 3 & 4 \end{bmatrix}$$

$$\begin{aligned} |A| &= -96 - 36 + 0 - (0 + 9 + 0) \\ &= -141 \end{aligned}$$

Alternatively, expanding in terms of the first row,

$$\begin{aligned} |A| &= 3 \begin{vmatrix} -8 & 1 \\ 3 & 4 \end{vmatrix} + 0 - 2 \begin{vmatrix} 6 & -8 \\ 0 & 3 \end{vmatrix} \\ &= 3(-32 - 3) - 2(18 - 0) \\ &= -105 - 36 = -141 \end{aligned}$$

Alternatively, expanding in terms of the second column,

$$\begin{aligned} |\mathbf{A}| &= 0 - 8 \begin{vmatrix} 3 & -2 \\ 0 & 4 \end{vmatrix} - 3 \begin{vmatrix} 3 & -2 \\ 6 & 1 \end{vmatrix} \\ &= 0 - 8(12 - 0) - 3(3 + 12) \\ &= -96 - 45 = -141 \end{aligned}$$

Note that cofactors are essentially determinants and can be evaluated by further expansion in terms of determinants of lower order. By repeated expansion, an n th-order determinant can be written in terms of second- or third-order determinants which can be evaluated easily. To simplify computation, a determinant should be expanded in terms of the row or column having the largest number of zero elements.

PROPERTIES OF DETERMINANTS

The following properties of determinants are frequently useful in their evaluation:

1. Interchanging the corresponding rows and columns of a determinant does not change its value, that is, $|\mathbf{A}| = |\mathbf{A}'|$.
2. If every element of a row (or of a column) of a determinant is zero, the value of the determinant is zero.
3. If every element of a row (or of a column) of a determinant is multiplied by the same constant, the value of the determinant is multiplied by that constant.
4. If two rows (or two columns) of a determinant are interchanged, the sign of the determinant is changed, but its absolute value is unchanged.
5. If two rows (or two columns) of a determinant are identical, the value of the determinant is zero.
6. The value of a determinant is not changed if each element of any row (or of any column), or each element multiplied by the same constant, is added to or subtracted from the corresponding element of any other row (or column).
7. The determinant of the product of two matrices is equal to the product of the determinants of the two matrices, that is, $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$.
8. The determinant of a diagonal matrix is equal to the product of its diagonal elements.

EXAMPLES

$$\text{If } \mathbf{A} = \begin{bmatrix} 1 & 3 & -3 \\ 2 & 0 & 1 \\ -1 & 4 & -2 \end{bmatrix},$$

(Property 1)

$$\begin{aligned} |\mathbf{A}| &= \begin{vmatrix} 1 & 3 & -3 \\ 2 & 0 & 1 \\ -1 & 4 & -2 \end{vmatrix} \\ &= 0 - 24 - 3 - (0 + 4 - 12) \\ &= -19 \end{aligned}$$

$$\begin{aligned}
 |\mathbf{A}'| &= \begin{vmatrix} 1 & 2 & -1 \\ 3 & 0 & 4 \\ -3 & 1 & -2 \end{vmatrix} \\
 &= 0 - 3 - 24 - (0 + 4 - 12) \\
 &= -19
 \end{aligned}$$

(Property 3)

$$\begin{aligned}
 |\mathbf{A}^*| &= \begin{vmatrix} 3 & 3 & -3 \\ 6 & 0 & 1 \\ -3 & 4 & -2 \end{vmatrix} \\
 &= 0 - 9 - 72 - (0 + 12 - 36) \\
 &= -57 \\
 &= 3|\mathbf{A}|
 \end{aligned}$$

(Note: First column of \mathbf{A} is multiplied by 3 to form \mathbf{A}^* .)

(Property 4)

$$\begin{aligned}
 |\mathbf{B}| &= \begin{vmatrix} 1 & 3 & -3 \\ -1 & 4 & -2 \\ 2 & 0 & 1 \end{vmatrix} = 4 + 0 - 12 - (-24 + 0 - 3) \\
 &= 19 \\
 &= -|\mathbf{A}|
 \end{aligned}$$

(Note: Second and third rows of \mathbf{A} are interchanged to form \mathbf{B} .)

(Property 6)

$$\begin{aligned}
 |\mathbf{C}| &= \begin{vmatrix} 1 & 3 & -3 \\ 3 & 4 & 0 \\ 2 & 0 & 1 \end{vmatrix} = 4 + 0 + 0 - (-24 + 0 + 9) \\
 &= 19 \\
 &= |\mathbf{B}|
 \end{aligned}$$

(Note: Second row of \mathbf{C} is equal to second row of \mathbf{B} plus two times the third row of \mathbf{B} .)

(Property 2)

$$\text{If } \mathbf{A} = \begin{bmatrix} 1 & 3 & -3 \\ 0 & 0 & 0 \\ -1 & 4 & -2 \end{bmatrix}, \text{ then}$$

$$\begin{aligned}
 |\mathbf{A}| &= \begin{vmatrix} 1 & 3 & -3 \\ 0 & 0 & 0 \\ -1 & 4 & 2 \end{vmatrix} \\
 &= 0 + 0 + 0 - (0 + 0 + 0) \\
 &= 0
 \end{aligned}$$

(Property 5)

$$\text{If } \mathbf{A} = \begin{bmatrix} 1 & 1 & -3 \\ 2 & 2 & 1 \\ -1 & -1 & -2 \end{bmatrix} \text{ then}$$

$$\begin{aligned}
 |\mathbf{A}| &= \begin{vmatrix} 1 & 1 & -3 \\ 2 & 2 & 1 \\ -1 & -1 & -2 \end{vmatrix} \\
 &= -4 + 6 - 1 - (6 - 1 - 4) \\
 &= 0
 \end{aligned}$$

(Property 7)

$$\text{If } \mathbf{A} = \begin{bmatrix} 1 & 3 & -3 \\ 2 & 0 & 1 \\ -1 & 4 & -2 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} -1 & 0 & 2 \\ 3 & -2 & -2 \\ 1 & 0 & -1 \end{bmatrix}, \text{ then}$$

$$\begin{aligned}
 |\mathbf{AB}| &= \begin{vmatrix} 5 & -6 & -1 \\ -1 & 0 & 3 \\ 11 & -8 & -8 \end{vmatrix} \\
 &= 0 - 8 - 198 - (0 - 120 - 48) \\
 &= -38
 \end{aligned}$$

$$|\mathbf{A}| = \begin{vmatrix} 1 & 3 & -3 \\ 2 & 0 & 1 \\ -1 & 4 & -2 \end{vmatrix} = -19$$

$$\begin{aligned}
 |\mathbf{B}| &= \begin{vmatrix} -1 & 0 & 2 \\ 3 & -2 & -2 \\ 1 & 0 & -1 \end{vmatrix} \\
 &= -2 + 0 + 0 - (-4 + 0 + 0) \\
 &= 2
 \end{aligned}$$

$$\text{and } |\mathbf{AB}| = |\mathbf{A}| |\mathbf{B}|.$$

(Property 8)

$$\text{If } \mathbf{A} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \text{ then}$$

$$\begin{aligned}
 |\mathbf{A}| &= \begin{vmatrix} 3 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{vmatrix} \\
 &= -6 + 0 + 0 - (0 + 0 + 0) \\
 &= -6
 \end{aligned}$$

7.8 The Inverse of a Matrix

The inverse of a matrix is used in solving simultaneous linear equations and in other applications. In this section the inverse of a matrix is defined and several methods of inversion are discussed; these methods include inversion by Gaussian elimination, inversion using adjoints and determinants, and inversion of partitioned matrices. Properties of inverses are also considered.

If for an $n \times n$ (square) matrix \mathbf{A} there is another $n \times n$ (square) matrix \mathbf{B} such that their product is the identity matrix of size n , that is, if

$$\mathbf{A}_{n \times n} \mathbf{B}_{n \times n} = \mathbf{I}_n = \mathbf{B}_{n \times n} \mathbf{A}_{n \times n}$$

then \mathbf{B} is said to be the *reciprocal* or *inverse* of \mathbf{A} and is written

$$\mathbf{B} = \mathbf{A}^{-1} = (a_{ij})^{-1} = (a^{ij}) = (b_{ij})$$

(it can be shown that $\mathbf{AB} = \mathbf{I} \Leftrightarrow \mathbf{BA} = \mathbf{I}$).

NOTE: In ordinary algebra, division of a quantity x by a quantity y is equivalent to multiplication of x by the reciprocal of y —that is,

$$\frac{x}{y} = x \left(\frac{1}{y} \right) = xy^{-1}$$

Finding the reciprocal or inverse of a matrix is an operation analogous to division in ordinary algebra; however, although every nonzero number has a reciprocal, there are square matrices, in addition to the null matrix, which do not have inverses. A matrix that has an inverse is said to be *nonsingular*; a matrix that has no inverse is said to be *singular*.

A matrix which is not square may have a left or right inverse. An $n \times s$ matrix \mathbf{A} is said to have \mathbf{B} as a *left inverse* if $\mathbf{BA} = \mathbf{I}$; in this case \mathbf{I} must be $s \times s$ and \mathbf{B} must be $s \times n$. Similarly, if $\mathbf{AC} = \mathbf{I}$, \mathbf{C} is said to be a *right inverse* of \mathbf{A} ; in this case \mathbf{I} must be $n \times n$ and \mathbf{C} must be $s \times n$. It can be shown that if \mathbf{A} has both a left inverse \mathbf{B} and a right inverse \mathbf{C} , then $\mathbf{B} = \mathbf{C} = \mathbf{A}^{-1}$ and \mathbf{A} is square (and nonsingular).

Obtaining the inverse of a matrix, referred to as *matrix inversion*, is essential for the matrix solution of sets of simultaneous linear equations; such equations arise, for example, in the solution of economic models and in linear programming problems.

Any set of simultaneous linear equations can be written in matrix notation; for example,

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n = c_1$$

$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n = c_2$$

$$\vdots$$

$$a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n = c_m$$

can be written

$$\mathbf{A}_{m \times n} \mathbf{x}_{n \times 1} = \mathbf{c}_{m \times 1}$$

If $m = n$ and \mathbf{A} has an inverse, that is, if \mathbf{A} is square and nonsingular, then the set of simultaneous linear equations represented by

$$\mathbf{A}_{n \times n} \mathbf{x}_{n \times 1} = \mathbf{c}_{n \times 1}$$

has the solution

$$\mathbf{x}_{n \times 1} = \mathbf{A}_{n \times n}^{-1} \mathbf{c}_{n \times 1}$$

Solution of simultaneous linear equations is considered in detail in section 7.9.

INVERSION OF 2×2 MATRICES

If $n = 2$, that is, if the set of simultaneous equations consists of two equations in two unknowns,

$$a_{11}x_1 + a_{12}x_2 = c_1$$

$$a_{21}x_1 + a_{22}x_2 = c_2$$

or, in matrix notation,

$$\mathbf{A}_{2 \times 2} \mathbf{x}_{2 \times 1} = \mathbf{c}_{2 \times 1}$$

then $\mathbf{A}_{2 \times 2}^{-1}$ can be obtained directly from its definition, as follows. If

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

and \mathbf{A}^{-1} is denoted by

$$\mathbf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

then, by definition, $\mathbf{AB} = \mathbf{I}$, that is,

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and thus

$$\begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and

$$a_{11}b_{11} + a_{12}b_{21} = 1$$

$$a_{11}b_{12} + a_{12}b_{22} = 0$$

$$a_{21}b_{11} + a_{22}b_{21} = 0$$

$$a_{21}b_{12} + a_{22}b_{22} = 1$$

Solving these four equations for the four b_{ij} 's,

$$b_{11} = \frac{a_{22}}{a_{11}a_{22} - a_{21}a_{12}}$$

$$b_{12} = \frac{-a_{12}}{a_{11}a_{22} - a_{21}a_{12}}$$

$$b_{21} = \frac{-a_{21}}{a_{11}a_{22} - a_{21}a_{12}}$$

$$b_{22} = \frac{a_{11}}{a_{11}a_{22} - a_{21}a_{12}}$$

Note that the denominator of each of these expressions is the *determinant* of \mathbf{A}

$$\det \mathbf{A} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}$$

If $\det \mathbf{A} = 0$, the b_{ij} 's are not defined and \mathbf{A}^{-1} cannot be obtained. It can be shown that a square matrix of any size has an inverse if and only if its determinant is nonzero.

EXAMPLE

Determine, if it exists, the inverse of the matrix

$$\begin{bmatrix} -1 & 6 \\ 4 & 3 \end{bmatrix}$$

$$\begin{vmatrix} -1 & 6 \\ 4 & 3 \end{vmatrix} = -3 - 24 = -27, \text{ so the inverse exists.}$$

$$b_{11} = \frac{3}{-27} = -\frac{1}{9}$$

$$b_{12} = \frac{-6}{-27} = \frac{2}{9}$$

$$b_{21} = \frac{-4}{-27} = \frac{4}{27}$$

$$b_{22} = \frac{-1}{-27} = \frac{1}{27}$$

and thus

$$\begin{bmatrix} -1 & 6 \\ 4 & 3 \end{bmatrix}^{-1} = \begin{bmatrix} -\frac{1}{9} & \frac{2}{9} \\ \frac{4}{27} & \frac{1}{27} \end{bmatrix}$$

Note that

$$\begin{bmatrix} -1 & 6 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} -\frac{1}{9} & \frac{2}{9} \\ \frac{4}{27} & \frac{1}{27} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

EXAMPLE

Determine, if it exists, the inverse of the matrix

$$\begin{bmatrix} -3 & 15 \\ 1 & -5 \end{bmatrix}$$

$$\begin{vmatrix} -3 & 15 \\ 1 & -5 \end{vmatrix} = 15 - 15 = 0, \text{ so the inverse does not exist, that is, the matrix is singular.}$$

PROBLEMS

For each of the following matrices, evaluate the determinant and find the inverse, if it exists.

Use partitioned matrices to check the results.

1. $\begin{bmatrix} 1 & 6 \\ 0 & 1 \end{bmatrix}$

7. $\begin{bmatrix} -1 & 3 \\ 2 & 7 \end{bmatrix}$

2. $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

8. $\begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix}$

3. $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

9. $\begin{bmatrix} 0 & -2 & -1 \\ 1 & -3 & 4 \\ -1 & -1 & -1 \end{bmatrix}$

4. $\begin{bmatrix} 0 & 1 \\ 1 & 6 \end{bmatrix}$

10. $\begin{bmatrix} 3 & 1 & 3 \\ 3 & 3 & 1 \\ 2 & 0 & 3 \end{bmatrix}$

5. $\begin{bmatrix} 2 & -3 \\ 4 & -6 \end{bmatrix}$

11. $\begin{bmatrix} 3 & 2 & -1 \\ 4 & 3 & -1 \\ -1 & 2 & 4 \end{bmatrix}$

6. $\begin{bmatrix} 2 & 2 \\ -4 & -4 \end{bmatrix}$

$$12. \begin{bmatrix} 1 & 2 & -1 \\ 0 & -3 & 2 \\ 4 & 1 & 0 \end{bmatrix}$$

$$14. \begin{bmatrix} 2 & 3 & 4 \\ 4 & 3 & 1 \\ 1 & 2 & 4 \end{bmatrix}$$

$$13. \begin{bmatrix} 2 & 4 & 6 \\ -1 & 2 & 3 \\ 1 & 4 & 9 \end{bmatrix}$$

Answers to Odd-Numbered Problems

$$1. \det = 1, \begin{bmatrix} 1 & -6 \\ 0 & 1 \end{bmatrix}$$

$$11. \det = 1, \begin{bmatrix} 14 & -10 & 1 \\ -15 & 11 & -1 \\ 11 & -8 & 1 \end{bmatrix}$$

$$3. \det = -1, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$13. \det = 24, \begin{bmatrix} \frac{1}{4} & -\frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{4} & -\frac{1}{6} & \frac{1}{3} \end{bmatrix}$$

$$5. \det = 0, \text{ so no inverse}$$

$$7. \det = -13, \begin{bmatrix} -\frac{7}{13} & \frac{3}{13} \\ \frac{2}{13} & \frac{1}{13} \end{bmatrix}$$

$$9. \det = 10, \begin{bmatrix} \frac{7}{10} & -\frac{1}{10} & -\frac{11}{10} \\ -\frac{3}{10} & -\frac{1}{10} & -\frac{1}{10} \\ -\frac{2}{5} & \frac{1}{5} & \frac{1}{5} \end{bmatrix}$$

INVERSION OF LARGER MATRICES

In principle, the inverse of any nonsingular matrix can be obtained from its definition in the same way as the inverse of a 2×2 matrix is obtained above. That is, if

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

and $\mathbf{B} = \mathbf{A}^{-1}$, then $\mathbf{AB} = \mathbf{I}$.

$$\begin{aligned} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix} &= \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix} \\ &= \begin{bmatrix} 1 & & & 0 \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \end{bmatrix} = \mathbf{I}_n \end{aligned}$$

where $c_{ik} = \sum_{j=1}^n a_{ij}b_{jk}$ and the n^2 elements of the inverse matrix (b_{jk}) can be obtained by solution of the n^2 equations

$$\sum_{j=1}^n a_{ij}b_{jk} = 1 \quad \text{if } i = k$$

$$\sum_{j=1}^n a_{ij}b_{jk} = 0 \quad \text{if } i \neq k$$

where $i = 1, 2, \dots, n$ and $k = 1, 2, \dots, n$.

For a matrix of size 3×3 or larger, this method of inversion is tedious. Fortunately, numerous alternative computational procedures for inverting a matrix are

available and inversion is also a standard program for most electronic computers. Two methods of inversion are discussed in following sections; both of these are feasible unless the matrix is quite large or its elements consist of several digits.

INVERSION USING ROW OR COLUMN OPERATIONS

To solve a system of simultaneous linear equations, certain simple operations are commonly used to convert the original system to an equivalent system—that is, to a system having the same solution as the original system—which is easier to solve. An equivalent system is obtained if (1) two equations are interchanged, (2) an equation is multiplied by any nonzero constant, and (3) the i th equation is replaced by the sum of the i th equation and k times the j th equation, where k is any nonzero constant.

Similarly, three *elementary row operations* on matrices are defined by (1) interchange of two rows, (2) multiplication of a row by any nonzero scalar, and (3) replacement of the i th row by the sum of the i th row and k times the j th row, where k is any nonzero scalar.

The corresponding operations on the columns of a matrix are *elementary column operations*.

THEOREM: If a matrix \mathbf{A} is converted to the identity matrix by a series of row operations or by a series of column operations, then the same series of operations performed on the identity matrix will convert it to \mathbf{A}^{-1} .

There is a standard method, invariably appropriate, for converting a nonsingular matrix to the identity matrix or for establishing the singularity of a matrix which cannot be converted to the identity matrix. This standard method is outlined and illustrated below. The generality of the method, which is appropriate for any matrix and which can be programmed for machine computation, makes it especially useful, even though for particular cases shorter methods are available.

Steps in Converting a Square Matrix to the Identity Matrix (Gaussian Elimination)

1. Divide the first row of the matrix by the element in its first column; use the resulting row to obtain zeros in the first column of each of the other rows.
2. Divide the second row by the element in its second column; use the resulting row to obtain zeros in the second column of each of the other rows.
- ⋮
- n . Divide the n th row by the element in its n th column; use the resulting row to obtain zeros in the n th column of each of the other rows.

NOTE: In the above procedure, if at the k th step the element in the k th column is zero, interchange that row with a subsequent row having a nonzero element in the k th column. Then proceed with the k th step.

To find the inverse of a matrix \mathbf{A} , it is customary to work with a tableau of the form

$$[\mathbf{A} | \mathbf{I}]$$

and change it by the standard procedure given above to the tableau

$$[\mathbf{I} | \mathbf{B}]$$

Then, by the theorem above, \mathbf{B} is \mathbf{A}^{-1} .

Even if A has no inverse, this procedure can be started. At some point a tableau will be obtained which is not of the appropriate form and which cannot be changed in the manner outlined above. It can be shown that if the above procedure cannot be completed to obtain $[I|B]$ from $[A|I]$, then A does not have an inverse.

EXAMPLE

Find the inverse, if it exists, of the matrix

$$\begin{bmatrix} 0 & -2 & -3 \\ 1 & 3 & 3 \\ -1 & -2 & -2 \end{bmatrix}$$

Following the standard procedure outlined above,

$$\left[\begin{array}{ccc|ccc} 0 & -2 & -3 & 1 & 0 & 0 \\ 1 & 3 & 3 & 0 & 1 & 0 \\ -1 & -2 & -2 & 0 & 0 & 1 \end{array} \right]$$

$$\left[\begin{array}{ccc|ccc} 1 & 3 & 3 & 0 & 1 & 0 \\ 0 & -2 & -3 & 1 & 0 & 0 \\ -1 & -2 & -2 & 0 & 0 & 1 \end{array} \right]$$

interchanging
first and
second rows

$$\left[\begin{array}{ccc|ccc} 1 & 3 & 3 & 0 & 1 & 0 \\ 0 & -2 & -3 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 \end{array} \right]$$

Step 1

$$\left[\begin{array}{ccc|ccc} 1 & 3 & 3 & 0 & 1 & 0 \\ 0 & 1 & \frac{3}{2} & -\frac{1}{2} & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 \end{array} \right]$$

Step 2

$$\left[\begin{array}{ccc|ccc} 1 & 0 & -\frac{3}{2} & \frac{3}{2} & 1 & 0 \\ 0 & 1 & \frac{3}{2} & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & \frac{1}{2} & 1 & 1 \end{array} \right]$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & -\frac{3}{2} & \frac{3}{2} & 1 & 0 \\ 0 & 1 & \frac{3}{2} & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & -1 & -2 & -2 \end{array} \right]$$

Step 3

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & -2 & -3 \\ 0 & 1 & 0 & 1 & 3 & 3 \\ 0 & 0 & 1 & -1 & -2 & -2 \end{array} \right]$$

Thus

$$\begin{bmatrix} 0 & -2 & -3 \\ 1 & 3 & 3 \\ -1 & -2 & -2 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & -2 & -3 \\ 1 & 3 & 3 \\ -1 & -2 & -2 \end{bmatrix}$$

and this matrix is equal to its inverse. Note that

$$\begin{bmatrix} 0 & -2 & -3 \\ 1 & 3 & 3 \\ -1 & -2 & -2 \end{bmatrix} \begin{bmatrix} 0 & -2 & -3 \\ 1 & 3 & 3 \\ -1 & -2 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

EXAMPLE

Find the inverse, if it exists, of the matrix

$$\begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 4 \\ 0 & 2 & 2 \end{bmatrix}$$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ -1 & 0 & 4 & 0 & 1 & 0 \\ 0 & 2 & 2 & 0 & 0 & 1 \end{array} \right]$$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 2 & 7 & 1 & 1 & 0 \\ 0 & 2 & 2 & 0 & 0 & 1 \end{array} \right]$$

Step 1

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & \frac{7}{2} & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 2 & 2 & 0 & 0 & 1 \end{array} \right]$$

Step 2

$$\left[\begin{array}{ccc|ccc} 1 & 0 & -4 & 0 & -1 & 0 \\ 0 & 1 & \frac{7}{2} & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & -5 & -1 & -1 & 1 \end{array} \right]$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & -4 & 0 & -1 & 0 \\ 0 & 1 & \frac{7}{2} & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 & \frac{1}{5} & \frac{1}{5} & -\frac{1}{5} \end{array} \right]$$

Step 3

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{4}{5} & -\frac{1}{5} & -\frac{4}{5} \\ 0 & 1 & 0 & -\frac{1}{5} & -\frac{1}{5} & \frac{7}{10} \\ 0 & 0 & 1 & \frac{1}{5} & \frac{1}{5} & -\frac{1}{5} \end{array} \right]$$

Thus

$$\begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 4 \\ 0 & 2 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{4}{5} & -\frac{1}{5} & -\frac{4}{5} \\ -\frac{1}{5} & -\frac{1}{5} & \frac{7}{10} \\ \frac{1}{5} & \frac{1}{5} & -\frac{1}{5} \end{bmatrix}$$

Note that

$$\begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 4 \\ 0 & 2 & 2 \end{bmatrix} \begin{bmatrix} \frac{4}{5} & -\frac{1}{5} & -\frac{4}{5} \\ -\frac{1}{5} & -\frac{1}{5} & \frac{7}{10} \\ \frac{1}{5} & \frac{1}{5} & -\frac{1}{5} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and

$$\begin{bmatrix} \frac{4}{5} & -\frac{1}{5} & -\frac{4}{5} \\ -\frac{1}{5} & -\frac{1}{5} & \frac{7}{10} \\ \frac{1}{5} & \frac{1}{5} & -\frac{1}{5} \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 4 \\ 0 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

EXAMPLE

Find the inverse, if it exists, of the matrix

$$\begin{bmatrix} 1 & 2 & -1 \\ -3 & 4 & 5 \\ -4 & 2 & 6 \end{bmatrix}$$

Following the standard procedure outlined above,

$$\left[\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ -3 & 4 & 5 & 0 & 1 & 0 \\ -4 & 2 & 6 & 0 & 0 & 1 \end{array} \right]$$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 10 & 2 & 3 & 1 & 0 \\ 0 & 10 & 2 & 4 & 0 & 1 \end{array} \right]$$

Step 1

$$\left[\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & \frac{1}{3} & \frac{3}{10} & \frac{1}{10} & 0 \\ 0 & 10 & 2 & 4 & 0 & 1 \end{array} \right]$$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & \frac{1}{3} & \frac{3}{10} & \frac{1}{10} & 0 \\ 0 & 0 & 0 & 1 & -1 & 1 \end{array} \right]$$

Step 2

The procedure cannot be continued and thus the matrix has no inverse. Note that the procedure can be discontinued at the end of Step 1; at the end of Step 1 the left-hand matrix has two identical rows and thus the determinant of the matrix is zero (Property 5) and it has no inverse. The determinant of the original matrix is also zero (Property 6) and it has no inverse.

INVERSION USING ADJOINTS AND DETERMINANTS

An alternative method of matrix inversion involves obtaining the adjoint and determinant of the matrix to be inverted. If A is nonsingular, that is, if $|A| \neq 0$, then

$$A^{-1} = \frac{1}{|A|} \text{adj } A$$

EXAMPLE

Find the inverse, if it exists, of the matrix

$$A = \begin{bmatrix} 0 & -2 & -3 \\ 1 & 3 & 3 \\ -1 & -2 & -2 \end{bmatrix}$$

$$|A| = 0 + 6 + 6 - (9 + 0 + 4) = -1$$

Alternatively, expanding by elements of the first row,

$$\begin{aligned} |A| &= 0 + (-2)(-1)^{1+2} \begin{vmatrix} 1 & 3 \\ -1 & -2 \end{vmatrix} + (-3)(-1)^{1+3} \begin{vmatrix} 1 & 3 \\ -1 & -2 \end{vmatrix} \\ &= 0 + 2(-2 + 3) + (-3)(-2 + 3) = -1 \end{aligned}$$

or, expanding by elements of the second column,

$$\begin{aligned} |A| &= (-2)(-1)^{1+2} \begin{vmatrix} 1 & 3 \\ -1 & -2 \end{vmatrix} + (3)(-1)^{2+2} \begin{vmatrix} 0 & -3 \\ -1 & -2 \end{vmatrix} + (-2)(-1)^{3+2} \begin{vmatrix} 0 & -3 \\ 1 & 3 \end{vmatrix} \\ &= 2(-2 + 3) + 3(0 - 3) + 2(0 + 3) = -1 \end{aligned}$$

Obtaining the adjoint,

$$C_{11} = (-1)^{1+1} \begin{vmatrix} 3 & 3 \\ -2 & -2 \end{vmatrix} = -6 + 6 = 0$$

$$C_{12} = (-1)^{1+2} \begin{vmatrix} 1 & 3 \\ -1 & -2 \end{vmatrix} = -(-2 + 3) = -1$$

$$C_{13} = (-1)^{1+3} \begin{vmatrix} 1 & 3 \\ -1 & -2 \end{vmatrix} = -2 + 3 = 1$$

$$C_{21} = (-1)^{2+1} \begin{vmatrix} -2 & -3 \\ -2 & -2 \end{vmatrix} = -(4 - 6) = 2$$

$$C_{22} = (-1)^{2+2} \begin{vmatrix} 0 & -3 \\ -1 & -2 \end{vmatrix} = 0 - 3 = -3$$

$$C_{23} = (-1)^{2+3} \begin{vmatrix} 0 & -2 \\ -1 & -2 \end{vmatrix} = -(0 - 2) = 2$$

$$C_{31} = (-1)^{3+1} \begin{vmatrix} -2 & -3 \\ 3 & 3 \end{vmatrix} = -6 + 9 = 3$$

$$C_{32} = (-1)^{3+2} \begin{vmatrix} 0 & -3 \\ 1 & 3 \end{vmatrix} = -(0 + 3) = -3$$

$$C_{33} = (-1)^{3+3} \begin{vmatrix} 0 & -2 \\ 1 & 3 \end{vmatrix} = 0 + 2 = 2$$

Thus

$$\text{adj } \mathbf{A} = (C_{ij})' = \begin{bmatrix} 0 & 2 & 3 \\ -1 & -3 & -3 \\ 1 & 2 & 2 \end{bmatrix}$$

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \text{adj } \mathbf{A} = \begin{bmatrix} 0 & -2 & -3 \\ 1 & 3 & 3 \\ -1 & -2 & -2 \end{bmatrix}$$

(as obtained in the example above by the standard procedure for row operations).

EXAMPLE

Find the inverse, if it exists, of the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 4 \\ 0 & 2 & 2 \end{bmatrix}$$

$$|\mathbf{A}| = 0 - 6 + 0 - (0 + 8 - 4) = -10$$

Alternatively, expanding by elements of the first column,

$$\begin{aligned} |\mathbf{A}| &= (-1)^{1+1} \begin{vmatrix} 0 & 4 \\ 2 & 2 \end{vmatrix} - (-1)^{2+1} \begin{vmatrix} 2 & 3 \\ 2 & 2 \end{vmatrix} + 0 \\ &= (0 - 8) + (4 - 6) = -10 \end{aligned}$$

or, expanding by elements of the second row,

$$\begin{aligned} |\mathbf{A}| &= (-1)(-1)^{2+1} \begin{vmatrix} 2 & 3 \\ 2 & 2 \end{vmatrix} + 0 + 4(-1)^{2+3} \begin{vmatrix} 1 & 2 \\ 0 & 2 \end{vmatrix} \\ &= (4 - 6) + 0 + (-4)(2 - 0) = -10 \end{aligned}$$

Obtaining the adjoint,

$$C_{11} = (-1)^{1+1} \begin{vmatrix} 0 & 4 \\ 2 & 2 \end{vmatrix} = 0 - 8 = -8$$

$$C_{12} = (-1)^{1+2} \begin{vmatrix} -1 & 4 \\ 0 & 2 \end{vmatrix} = -(-2 + 0) = 2$$

$$C_{13} = (-1)^{1+3} \begin{vmatrix} -1 & 0 \\ 0 & 2 \end{vmatrix} = -2 + 0 = -2$$

$$C_{21} = (-1)^{2+1} \begin{vmatrix} 2 & 3 \\ 2 & 2 \end{vmatrix} = -(4 - 6) = 2$$

$$C_{22} = (-1)^{2+2} \begin{vmatrix} 1 & 3 \\ 0 & 2 \end{vmatrix} = 2 + 0 = 2$$

$$C_{23} = (-1)^{2+3} \begin{vmatrix} 1 & 2 \\ 0 & 2 \end{vmatrix} = -(2 + 0) = -2$$

$$C_{31} = (-1)^{3+1} \begin{vmatrix} 2 & 3 \\ 0 & 4 \end{vmatrix} = 8 + 0 = 8$$

$$C_{32} = (-1)^{3+2} \begin{vmatrix} 1 & 3 \\ -1 & 4 \end{vmatrix} = -(4 + 3) = -7$$

$$C_{33} = (-1)^{3+3} \begin{vmatrix} 1 & 2 \\ -1 & 0 \end{vmatrix} = 0 + 2 = 2$$

Thus

$$\text{adj } \mathbf{A} = (C_{ij})' = \begin{bmatrix} -8 & 2 & 8 \\ 2 & 2 & -7 \\ -2 & -2 & 2 \end{bmatrix}$$

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \text{adj } \mathbf{A} = \begin{bmatrix} \frac{4}{5} & -\frac{1}{5} & -\frac{4}{5} \\ -\frac{1}{5} & -\frac{1}{5} & \frac{7}{10} \\ \frac{1}{5} & \frac{1}{5} & -\frac{1}{5} \end{bmatrix}$$

(as obtained in the example above by the standard procedure for row operations).

EXAMPLE

Find the inverse, if it exists, of the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & -1 \\ -3 & 4 & 5 \\ -4 & 2 & 6 \end{bmatrix}$$

$$|\mathbf{A}| = 24 + 6 - 40 - (16 + 10 - 36) = 0$$

Thus \mathbf{A} is singular.

In general, inversion by row or column operations is less laborious if the elements of a matrix are small whole numbers; inversion by determinant and adjoint is less laborious if the elements of a matrix are large numbers or fractions. For large matrices both of these methods are tedious, and electronic computers should be used, if available.

Unless there is good reason to believe that a matrix is nonsingular, as in some practical applications, it is advisable first to find its determinant, even if inversion by row or column operations is intended. A number of row or column operations may be necessary before the singularity of a matrix becomes apparent using this method of inversion, and much useless computation is thus avoided by establishing the existence of an inverse before attempting (by any method) to compute it.

INVERSION OF PARTITIONED MATRICES

It is sometimes convenient to obtain the inverse of a matrix in partitioned form. If an $n \times n$ matrix A is partitioned

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

where A_{11} is $n_1 \times n_1$, A_{12} is $n_1 \times n_2$, A_{21} is $n_2 \times n_1$, A_{22} is $n_2 \times n_2$, and $n_1 + n_2 = n$, then

$$A^{-1} = \begin{bmatrix} A_{11}^{-1}[I + A_{12}B^{-1}A_{21}A_{11}^{-1}] & -A_{11}^{-1}A_{12}B^{-1} \\ -B^{-1}A_{21}A_{11}^{-1} & B^{-1} \end{bmatrix}$$

where $B = A_{22} - A_{21}A_{11}^{-1}A_{12}$ and A_{11} and B are nonsingular. This result can be verified by multiplication:

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} A_{11}^{-1}[I + A_{12}B^{-1}A_{21}A_{11}^{-1}] & -A_{11}^{-1}A_{12}B^{-1} \\ -B^{-1}A_{21}A_{11}^{-1} & B^{-1} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

since

$$\begin{aligned} & A_{11}A_{11}^{-1}[I + A_{12}B^{-1}A_{21}A_{11}^{-1}] - A_{12}B^{-1}A_{21}A_{11}^{-1} = I \\ & -A_{11}A_{11}^{-1}A_{12}B^{-1} + A_{12}B^{-1} = 0 \\ & A_{21}A_{11}^{-1}[I + A_{12}B^{-1}A_{21}A_{11}^{-1}] - A_{22}B^{-1}A_{21}A_{11}^{-1} \\ & \quad = [I + \underbrace{A_{21}A_{11}^{-1}A_{12}B^{-1} - A_{22}B^{-1}}_{[A_{21}A_{11}^{-1}A_{12} - A_{22}]B^{-1}}]A_{21}A_{11}^{-1} \\ & \quad = [I - BB^{-1}]A_{21}A_{11}^{-1} \\ & \quad = 0 \\ & -A_{21}A_{11}^{-1}A_{12}B^{-1} + A_{22}B^{-1} = [-A_{21}A_{11}^{-1}A_{12} + A_{22}]B^{-1} \\ & \quad = BB^{-1} \\ & \quad = I \end{aligned}$$

EXAMPLE

Find the inverse, if it exists, of the matrix

$$\begin{bmatrix} 0 & -2 & -3 \\ 1 & 3 & 3 \\ -1 & -2 & -2 \end{bmatrix}$$

Partitioning the matrix,

$$\begin{bmatrix} 0 & -2 & -3 \\ 1 & 3 & 3 \\ -1 & -2 & -2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

and

$$\begin{aligned}
 \mathbf{A}_{11}^{-1}[\mathbf{I} + \mathbf{A}_{12} \mathbf{B}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1}] &= \begin{bmatrix} \frac{3}{2} & 1 \\ -\frac{1}{2} & 0 \end{bmatrix} \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} -3 \\ 3 \end{bmatrix} [-2] [-1, -2] \begin{bmatrix} \frac{3}{2} & 1 \\ -\frac{1}{2} & 0 \end{bmatrix} \right) \\
 &= \begin{bmatrix} \frac{3}{2} & 1 \\ -\frac{1}{2} & 0 \end{bmatrix} \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} -6 & -12 \\ 6 & 12 \end{bmatrix} \begin{bmatrix} \frac{3}{2} & 1 \\ -\frac{1}{2} & 0 \end{bmatrix} \right) \\
 &= \begin{bmatrix} \frac{3}{2} & 1 \\ -\frac{1}{2} & 0 \end{bmatrix} \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} -3 & -6 \\ 3 & 6 \end{bmatrix} \right) \\
 &= \begin{bmatrix} \frac{3}{2} & 1 \\ -\frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} -2 & -6 \\ 3 & 7 \end{bmatrix} \\
 &= \begin{bmatrix} 0 & -2 \\ 1 & 3 \end{bmatrix}
 \end{aligned}$$

where

$$\begin{aligned}
 \mathbf{B} &= \mathbf{A}_{22} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12} = [-2] - [-1, -2] \begin{bmatrix} \frac{3}{2} & 1 \\ -\frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} -3 \\ 3 \end{bmatrix} \\
 &= [-2] - [-\frac{1}{2}, -1] \begin{bmatrix} -3 \\ 3 \end{bmatrix} \\
 &= -2 - (-\frac{3}{2}) \\
 &= -\frac{1}{2}
 \end{aligned}$$

and

$$\begin{aligned}
 \mathbf{B}^{-1} &= [-2] \\
 -\mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{B}^{-1} &= \begin{bmatrix} -\frac{3}{2} & -1 \\ \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} -3 \\ 3 \end{bmatrix} [-2] \\
 &= \begin{bmatrix} -\frac{3}{2} & -1 \\ \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} 6 \\ -6 \end{bmatrix} \\
 &= \begin{bmatrix} -3 \\ 3 \end{bmatrix} \\
 -\mathbf{B}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} &= [2] [-1, -2] \begin{bmatrix} \frac{3}{2} & 1 \\ -\frac{1}{2} & 0 \end{bmatrix} = [-1, -2] \\
 &= [-2, -4] \begin{bmatrix} \frac{3}{2} & 1 \\ -\frac{1}{2} & 0 \end{bmatrix} \\
 &= [-1, -2]
 \end{aligned}$$

Thus

$$\begin{bmatrix} 0 & -2 & -3 \\ 1 & 3 & 3 \\ -1 & -2 & -2 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & -2 & -3 \\ 1 & 3 & 3 \\ -1 & -2 & -2 \end{bmatrix}$$

as above.

EXAMPLE

Find the inverse, if it exists, of the matrix

$$\begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 4 \\ 0 & 2 & 2 \end{bmatrix}$$

Partitioning the matrix,

$$\left[\begin{array}{c|cc} 1 & 2 & 3 \\ \hline -1 & 0 & 4 \\ \hline 0 & 2 & 2 \end{array} \right] = \left[\begin{array}{c|c} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \hline \mathbf{A}_{21} & \mathbf{A}_{22} \end{array} \right]$$

and

$$\begin{aligned} \mathbf{A}_{11}^{-1}[\mathbf{I} + \mathbf{A}_{12} \mathbf{B}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1}] &= [1] \left([1] + [2, 3] \begin{bmatrix} -\frac{1}{5} & \frac{7}{10} \\ \frac{1}{5} & -\frac{1}{5} \end{bmatrix} \begin{bmatrix} -1 \\ 0 \end{bmatrix} [1] \right) \\ &= [1] \left([1] + [\frac{1}{5}, \frac{4}{5}] \begin{bmatrix} -1 \\ 0 \end{bmatrix} \right) \\ &= 1(1 - \frac{1}{5}) \\ &= \frac{4}{5} \end{aligned}$$

where

$$\begin{aligned} \mathbf{B} &= \mathbf{A}_{22} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12} = \begin{bmatrix} 0 & 4 \\ 2 & 2 \end{bmatrix} - \begin{bmatrix} -1 \\ 0 \end{bmatrix} [1] [2, 3] \\ &= \begin{bmatrix} 0 & 4 \\ 2 & 2 \end{bmatrix} - \begin{bmatrix} -2 & -3 \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 7 \\ 2 & 2 \end{bmatrix} \end{aligned}$$

and

$$\begin{aligned} \mathbf{B}^{-1} &= \begin{bmatrix} -\frac{1}{5} & \frac{7}{10} \\ \frac{1}{5} & -\frac{1}{5} \end{bmatrix} \\ -\mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{B}^{-1} &= [-1] [2, 3] \begin{bmatrix} -\frac{1}{5} & \frac{7}{10} \\ \frac{1}{5} & -\frac{1}{5} \end{bmatrix} \\ &= [-\frac{1}{5}, -\frac{4}{5}] \\ -\mathbf{B}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} &= \begin{bmatrix} \frac{1}{5} & -\frac{7}{10} \\ -\frac{1}{5} & \frac{1}{5} \end{bmatrix} \begin{bmatrix} -1 \\ 0 \end{bmatrix} [1] \\ &= \begin{bmatrix} -\frac{1}{5} \\ \frac{1}{5} \end{bmatrix} \end{aligned}$$

Thus

$$\left[\begin{array}{ccc} 1 & 2 & 3 \\ -1 & 0 & 4 \\ 0 & 2 & 2 \end{array} \right]^{-1} = \begin{bmatrix} \frac{4}{5} & -\frac{1}{5} & -\frac{4}{5} \\ -\frac{1}{5} & -\frac{1}{5} & \frac{7}{10} \\ \frac{1}{5} & \frac{1}{5} & -\frac{1}{5} \end{bmatrix}$$

as above.

PROPERTIES OF INVERSES

The following properties of inverses are frequently useful in their evaluation.

1. The inverse of the inverse of a matrix is the original matrix; that is, $[\mathbf{A}^{-1}]^{-1} = \mathbf{A}$.
2. The determinant of the inverse of a matrix is equal to the reciprocal of the determinant of the matrix; that is, $|\mathbf{A}^{-1}| = 1/|\mathbf{A}|$.
3. The inverse of the transpose of a matrix is equal to the transpose of the inverse of the matrix; that is, $[\mathbf{A}']^{-1} = [\mathbf{A}^{-1}]'$.
4. The inverse of the product of two matrices is equal to the product of their inverses in reverse order; that is, $[\mathbf{AB}]^{-1} = \mathbf{B}^{-1} \mathbf{A}^{-1}$.

EXAMPLES

$$\text{If } \mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 4 \\ 0 & 2 & 2 \end{bmatrix}, \text{ then } \mathbf{A}^{-1} = \begin{bmatrix} \frac{4}{3} & -\frac{1}{3} & -\frac{4}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{7}{10} \\ \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} \end{bmatrix}$$

(Property 1)

$$\begin{aligned} [\mathbf{A}^{-1}]^{-1} &= \begin{bmatrix} \frac{4}{3} & -\frac{1}{3} & -\frac{4}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{7}{10} \\ \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 4 \\ 0 & 2 & 2 \end{bmatrix} \\ &= \mathbf{A} \end{aligned}$$

Thus $[\mathbf{A}^{-1}]^{-1} = \mathbf{A}$.

(Property 2)

$$\begin{aligned} |\mathbf{A}| &= \begin{vmatrix} 1 & 2 & 3 \\ -1 & 0 & 4 \\ 0 & 2 & 2 \end{vmatrix} \\ &= 0 - 6 + 0 - (0 + 8 - 4) \\ &= -10 \\ |\mathbf{A}^{-1}| &= \begin{vmatrix} \frac{4}{3} & -\frac{1}{3} & -\frac{4}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{7}{10} \\ \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} \end{vmatrix} \\ &= \frac{1}{123} [4 + 4 - \frac{7}{2} - (4 + 14 - 1)] \\ &= (\frac{1}{123}) (-\frac{25}{2}) \\ &= -\frac{1}{10} \end{aligned}$$

$$\text{Thus } |\mathbf{A}^{-1}| = \frac{1}{|\mathbf{A}|}.$$

(Property 3)

$$\begin{aligned} (\mathbf{A}')^{-1} &= \begin{bmatrix} 1 & -1 & 0 \\ 2 & 0 & 2 \\ 3 & 4 & 2 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} \frac{4}{3} & -\frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \\ -\frac{4}{3} & \frac{7}{10} & -\frac{1}{3} \end{bmatrix} \\ [\mathbf{A}^{-1}]' &= \begin{bmatrix} \frac{4}{3} & -\frac{1}{3} & -\frac{4}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{7}{10} \\ \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} \end{bmatrix} \\ &= \begin{bmatrix} \frac{4}{3} & -\frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \\ -\frac{4}{3} & \frac{7}{10} & -\frac{1}{3} \end{bmatrix} \end{aligned}$$

Thus $[\mathbf{A}']^{-1} = [\mathbf{A}^{-1}]'$.

(Property 4)

$$\text{If } \mathbf{B} = \begin{bmatrix} -1 & 3 & 1 \\ 0 & 2 & 0 \\ -2 & 0 & 4 \end{bmatrix}, \text{ then } \mathbf{B}^{-1} = \begin{bmatrix} -2 & 3 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \\ -1 & \frac{3}{2} & \frac{1}{2} \end{bmatrix}$$

$$\mathbf{AB} = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 4 \\ 0 & 2 & 2 \end{bmatrix} \begin{bmatrix} -1 & 3 & 1 \\ 0 & 2 & 0 \\ -2 & 0 & 4 \end{bmatrix}$$

$$= \begin{bmatrix} -7 & 7 & 13 \\ -7 & -3 & 15 \\ -4 & 4 & 8 \end{bmatrix}$$

$$[\mathbf{AB}]^{-1} = \begin{bmatrix} -7 & 7 & 13 \\ -7 & -3 & 15 \\ -4 & 4 & 8 \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} -\frac{21}{10} & -\frac{1}{10} & \frac{18}{5} \\ -\frac{1}{10} & -\frac{1}{10} & \frac{7}{20} \\ -1 & 0 & \frac{7}{4} \end{bmatrix}$$

$$\mathbf{B}^{-1}\mathbf{A}^{-1} = \begin{bmatrix} -2 & 3 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \\ -1 & \frac{3}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{4}{5} & -\frac{1}{5} & -\frac{4}{5} \\ -\frac{1}{5} & -\frac{1}{5} & \frac{7}{10} \\ \frac{1}{5} & \frac{1}{5} & -\frac{1}{5} \end{bmatrix}$$

$$= \begin{bmatrix} -\frac{21}{10} & -\frac{1}{10} & \frac{18}{5} \\ -\frac{1}{10} & -\frac{1}{10} & \frac{7}{20} \\ -1 & 0 & \frac{7}{4} \end{bmatrix}$$

Thus $[\mathbf{AB}]^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.

PROBLEMS

1. If

$$\mathbf{A} = \begin{bmatrix} 3 & 2 & -2 \\ 0 & 1 & 4 \\ -1 & 0 & 5 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad \mathbf{D} = \begin{bmatrix} 3 & -2 \\ 2 & 3 \end{bmatrix}$$

find a. $\det [(\mathbf{AB})'\mathbf{C}' - \mathbf{D}]$ b. $[\mathbf{BC}]^{-1}$ c. $[\mathbf{D}^2]^{-1}$

2. If

$$\mathbf{A} = \begin{bmatrix} 3 & -1 \\ -4 & 0 \\ 2 & 1 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 2 & 1 \\ -1 & -2 \\ 1 & 1 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 3 & 4 \\ 2 & 2 \end{bmatrix}$$

find a. $[\mathbf{A} - \mathbf{BC}]'$ b. $[\mathbf{A}'\mathbf{B}]^{-1}$

3. If

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 1 & 2 \end{bmatrix} \quad \mathbf{D} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

find a. $\mathbf{A} + \mathbf{B} - 2\mathbf{DC}$ b. $\mathbf{A}^{-1}\mathbf{B}^{-1}$ c. $\mathbf{C}[\mathbf{A} + \mathbf{B}]\mathbf{B}^{-1}$ d. $\mathbf{CA}^{-1}\mathbf{B}^{-1}\mathbf{D}$

4. If

$$A = \begin{bmatrix} 0 & 1 & -1 \\ 3 & -2 & 3 \\ 2 & -2 & 3 \end{bmatrix} \quad B = \begin{bmatrix} 4 & -3 & 3 \\ 2 & -1 & 2 \\ -3 & 3 & -2 \end{bmatrix}$$

show that a. $A^2 = B^2 = \left[\frac{1}{2}(A + B)\right]^2 = I$
 b. $[A - B]^2 = O$

Answers to Odd-Numbered Problems

1. a. -39

b. inverse does not exist

$$c. \begin{bmatrix} \frac{5}{169} & \frac{12}{169} \\ -\frac{12}{169} & \frac{5}{169} \end{bmatrix}$$

$$3. a. \begin{bmatrix} -3 & -7 \\ -9 & -13 \end{bmatrix}$$

$$c. \begin{bmatrix} 5 & 7 \\ 13 & 15 \\ 5 & 7 \end{bmatrix}$$

$$b. \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

$$d. \begin{bmatrix} 1 & 1 & 1 \\ \frac{5}{3} & \frac{7}{3} & \frac{5}{3} \\ 1 & 1 & 1 \end{bmatrix}$$

7.9 Simultaneous Linear Equations

This section concerns simultaneous linear equation systems. Linear dependence of a set of vectors and the related concept of rank of a matrix are defined and applied to the problem of determining whether a set of simultaneous linear equations has a unique solution. Methods of obtaining the solution of a set of simultaneous linear equations are also discussed.

LINEAR DEPENDENCE AND RANK

A set of m vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m$ each with n elements is said to be *linearly dependent* if there is a nontrivial linear combination of the vectors that is equal to the zero vector with n elements. That is, if there is a set of numbers $\lambda_1, \lambda_2, \dots, \lambda_m$ (not all zero) such that

$$\lambda_1 \mathbf{a}_1 + \lambda_2 \mathbf{a}_2 + \dots + \lambda_m \mathbf{a}_m = \sum_{i=1}^m \lambda_i \mathbf{a}_i = \mathbf{0}$$

then the set of vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m$ is said to be *linearly dependent*. If there is no set of λ 's (except all zeros) such that $\sum_{i=1}^m \lambda_i \mathbf{a}_i = \mathbf{0}$, the set of vectors is said to be *linearly independent*.

A matrix can be thought of as a set of row vectors or a set of column vectors. It can be shown that, for any matrix, the number of linearly independent rows is equal to the number of linearly independent columns; this number is the *rank* of the matrix. Thus if a matrix is $m \times n$ and its rank is denoted by r , then $r \leq \min(m, n)$.

EXAMPLE

Determine the rank of the matrix

$$\begin{bmatrix} 1 & 4 \\ 5 & 10 \\ 3 & 2 \end{bmatrix}$$

Since $2(\text{first row}) + (\text{third row}) - (\text{second row}) = 0$, the three rows are linearly dependent. However, each pair of two rows is linearly independent; the two columns are also linearly independent. The rank of the matrix is 2.

EXAMPLE

Determine the rank of the matrix

$$\begin{bmatrix} 3 & 6 \\ 1 & 2 \\ 2 & 4 \end{bmatrix}$$

The three rows are linearly dependent, since

$$(\text{first row}) - (\text{second row}) - (\text{third row}) = 0$$

Each pair of two rows is linearly dependent, since

$$(\text{first row}) - 3(\text{second row}) = 0$$

$$2(\text{first row}) - 3(\text{third row}) = 0$$

$$2(\text{second row}) - (\text{third row}) = 0$$

Note that the two columns are also linearly dependent, since

$$2(\text{first column}) - (\text{second column}) = 0$$

The rank of the matrix is 1.

The following very useful result provides a systematic method of testing for linear dependence: Consider all the square submatrices of \mathbf{A} whose determinants are non-zero. The rank of \mathbf{A} is the order of the determinant of largest order. Thus one method of computing the rank of a matrix is to look for the largest nonzero determinant; the order of this determinant is the rank of the matrix.

The following properties of rank are useful in determining the rank of a matrix:

1. Since the determinant of a diagonal matrix is equal to the product of its diagonal elements, if \mathbf{A} is a diagonal matrix, then $r(\mathbf{A})$ is the number of nonzero diagonal elements in \mathbf{A} . In particular, $r(\mathbf{I}_n) = n$.
2. Since any submatrix of \mathbf{A}' is the transpose of a submatrix of \mathbf{A} and since $|\mathbf{B}| = |\mathbf{B}'|$, $r(\mathbf{A}') = r(\mathbf{A})$.
3. The rank of the product of two matrices cannot exceed the smaller rank of the two matrices; that is, $r(\mathbf{AB}) \leq \min\{r(\mathbf{A}), r(\mathbf{B})\}$.
4. If \mathbf{A} is an $n \times n$ (square) matrix, then $r(\mathbf{A}) = n$ if and only if \mathbf{A} is nonsingular; $r(\mathbf{A}) < n$ if and only if \mathbf{A} is singular. Thus if a square matrix is singular, its rows (and also its columns) are linearly dependent; if a square matrix is nonsingular, its rows (and also its columns) are linearly independent.

Consider the solution of a general system of n linear equations in n variables x_1, x_2, \dots, x_n ,

$$\mathbf{y} = \mathbf{Ax}$$

where \mathbf{y} is $n \times 1$, \mathbf{A} is $n \times n$, and \mathbf{x} is $n \times 1$. If \mathbf{A} is nonsingular, the unique solution is

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{y}$$

and, conversely, if $\mathbf{y} = \mathbf{Ax}$ has a unique solution, \mathbf{A} is nonsingular.

More generally, if $y = Ax$, then the whole set of equations is consistent and has at least one solution if $r(A) = r(A|y)$. The solution is unique if and only if $r(A) = r(A|y) = n$, that is, if and only if A is nonsingular.

In the special (homogeneous) case $y = 0$, if A is nonsingular, the unique solution is $x = 0$ and there cannot be a nonzero solution. Thus, when the set of equations $Ax = 0$ does have a nonzero solution, A must be singular.

Suppose that $y = Ax$ and A is $m \times n$ (not square), that is, there are m equations in n variables where m may be less than, equal to, or greater than n ; then $r(A) \leq \min(m, n)$.

The following summary concerning solution of simultaneous linear equations applies to the general case of m equations in n variables and also to the special case of n equations in n variables.

If $r(A|y) = r(A)$, the equations in the set are logically consistent and there is at least one solution

If $r(A|y) = r(A) = n$, there is a unique solution

If $r(A|y) = r(A) < n$, there are infinitely many solutions and the rows (and columns) of A are linearly dependent

If $r(A|y) \neq r(A)$, the equations in the set are not logically consistent and there is no solution

EXAMPLE

Consider the following set of simultaneous linear equations

$$x + 2y - z = 10$$

$$2x + 4y - 2z = 5$$

$$x + y + z = 6$$

The determinant of the coefficient matrix A is zero

$$\begin{vmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \\ 1 & 1 & 1 \end{vmatrix} = 4 - 2 - 4 - (-4 - 2 + 4) = 0$$

so the rank of the matrix is not 3; the rank is 2 since, for example,

$$\begin{vmatrix} 2 & 4 \\ 1 & 1 \end{vmatrix} = 2 - 4 = -2 \neq 0$$

The augmented coefficient matrix $[A|Y]$

$$\begin{bmatrix} 1 & 2 & -1 & 10 \\ 2 & 4 & -2 & 5 \\ 1 & 1 & 1 & 6 \end{bmatrix}$$

has rank 3 since, for example,

$$\begin{vmatrix} 2 & -1 & 10 \\ 4 & -2 & 5 \\ 1 & 1 & 6 \end{vmatrix} = -24 + 40 - 5 - (-20 + 10 - 24) = 45 \neq 0$$

Thus $r(A) \neq r(A|y)$ and the set of equations has no solution since they are not all consistent. Note that the first and second equations are clearly inconsistent.

EXAMPLE

Consider the following set of simultaneous linear equations

$$x + 2y - z = 10$$

$$2x + 4y - 2z = 20$$

$$x + y + z = 6$$

The coefficient matrix is the same as that for the preceding example and its rank is 2. The augmented coefficient matrix $(A; y)$

$$\begin{bmatrix} 1 & 2 & -1 & 10 \\ 2 & 4 & -2 & 20 \\ 1 & 1 & 1 & 6 \end{bmatrix}$$

is also rank 2, since

$$\begin{vmatrix} 1 & 2 & 10 \\ 2 & 4 & 20 \\ 1 & 1 & 6 \end{vmatrix} = \begin{vmatrix} 1 & -1 & 10 \\ 2 & -2 & 20 \\ 1 & 1 & 6 \end{vmatrix} = \begin{vmatrix} 2 & -1 & 10 \\ 4 & -2 & 20 \\ 1 & 1 & 6 \end{vmatrix} = 0$$

Thus $r(A) = r(A; y) = 2 \neq n = 3$, and the set of equations has infinitely many solutions; the rows (and columns) of the coefficient matrix are linearly dependent. Note that the first and second equations are linearly dependent.

EXAMPLE

Consider the following set of simultaneous linear equations

$$x + 2y - z = 10$$

$$2x - 4y - 2z = 5$$

$$x + y + z = 6$$

The determinant of the coefficient matrix A is nonzero

$$\begin{vmatrix} 1 & 2 & -1 \\ 2 & -4 & -2 \\ 1 & 1 & 1 \end{vmatrix} = -4 - 2 - 4 - (4 - 2 + 4) = -16$$

and the rank of the matrix is 3; the rank of the augmented coefficient matrix is also 3. Thus $r(A) = r(A; y) = 3 = n$ and the set of equations has a unique solution.

PROBLEMS

Determine the rank of each of the following matrices.

1. $\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$

4. $\begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 2 & -2 & 0 \end{bmatrix}$

2. $\begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 1 \\ 0 & 4 & 8 \end{bmatrix}$

5. $\begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$

3. $\begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & 3 \\ 0 & 1 & 2 \end{bmatrix}$

6. $\begin{bmatrix} -2 & 1 & 4 \\ 0 & -1 & 5 \\ -2 & 0 & -9 \end{bmatrix}$

$$7. \begin{bmatrix} 3 & 1 & -2 \\ 2 & 5 & 4 \\ -4 & 3 & 1 \end{bmatrix}$$

$$8. \begin{bmatrix} -1 & 2 & -3 \\ 2 & -3 & 1 \\ 0 & 1 & -5 \end{bmatrix}$$

$$9. \begin{bmatrix} 3 & -1 & 0 & 2 \\ -2 & 1 & 5 & -2 \\ 0 & -4 & 6 & -3 \\ -3 & 5 & -6 & 1 \end{bmatrix}$$

$$10. \begin{bmatrix} 2 & -1 & 3 & 0 \\ -3 & 3 & 4 & 2 \\ 4 & -5 & -11 & -4 \\ -1 & 2 & 7 & 2 \end{bmatrix}$$

Answers to Odd-Numbered Problems

1. rank 2

3. rank 3

5. rank 1

7. rank 3

9. rank 3

SOLUTION OF SIMULTANEOUS LINEAR EQUATIONS

There are various methods for obtaining the solution of a set of simultaneous linear equations which have a unique solution. Several of these methods are discussed in this section. If one of the methods is applied to a set of linear equations not having a unique solution, the method cannot be completed, thus indicating that there is no unique solution.

As noted above, a set of n simultaneous linear equations in n unknowns can be written in matrix notation as

$$\mathbf{A}_{n \times n} \mathbf{x}_{n \times 1} = \mathbf{c}_{n \times 1}$$

and the solution obtained by inversion of \mathbf{A} can be denoted by

$$\mathbf{x}_{n \times 1} = \mathbf{A}_{n \times n}^{-1} \mathbf{c}_{n \times 1}$$

Alternatively, the standard procedure described previously can be used to change the tableau

$$(\mathbf{A} | \mathbf{I} | \mathbf{c})$$

into the tableau

$$(\mathbf{I} | \mathbf{A}^{-1} | \mathbf{x})$$

from which the solution can be read directly.

A third matrix method for the solution of simultaneous linear equations is given by *Cramer's rule*: The solution of

$$\mathbf{A}_{n \times n} \mathbf{x}_{n \times 1} = \mathbf{c}_{n \times 1}$$

can be obtained as the ratios of determinants

$$\begin{aligned}
 x_1 &= \frac{\begin{vmatrix} c_1 & a_{12} & \cdots & a_{1n} \\ c_2 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_n & a_{n2} & \cdots & a_{nn} \end{vmatrix}}{|A|} \\
 x_2 &= \frac{\begin{vmatrix} a_{11} & c_1 & \cdots & a_{1n} \\ a_{21} & c_2 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & c_n & \cdots & a_{nn} \end{vmatrix}}{|A|} \\
 &\vdots \\
 x_n &= \frac{\begin{vmatrix} a_{11} & a_{12} & \cdots & c_1 \\ a_{21} & a_{22} & \cdots & c_2 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & c_n \end{vmatrix}}{|A|}
 \end{aligned}$$

For each x_i , $i = 1, 2, \dots, n$, the denominator is the determinant of the coefficient matrix and the numerator is the determinant of the coefficient matrix with the i th column replaced by the column of constant terms from the right-hand side of the equations. If there is no unique solution for a set of linear equations, $|A| = 0$ and these quotients are not defined.

EXAMPLE

Solve the set of simultaneous linear equations

$$3x_1 + x_2 - x_3 = 2$$

$$x_1 - 2x_2 + x_3 = -9$$

$$4x_1 + 3x_2 + 2x_3 = 1$$

Apply the standard procedure to the tableau

$$(A|I|c)$$

to obtain the tableau

$$(I|A^{-1}|x)$$

Then the solution can be read directly from the tableau

$$(I|x)$$

or obtained as

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{c}$$

$$\left[\begin{array}{ccc|ccc|c} 3 & 1 & -1 & 1 & 0 & 0 & 2 \\ 1 & -2 & 1 & 0 & 1 & 0 & -9 \\ 4 & 3 & 2 & 0 & 0 & 1 & 1 \end{array} \right]$$

$$\left[\begin{array}{ccc|ccc|c} 1 & -2 & 1 & 0 & 1 & 0 & -9 \\ 3 & 1 & -1 & 1 & 0 & 0 & 2 \\ 4 & 3 & 2 & 0 & 0 & 1 & 1 \end{array} \right]$$

$$\left[\begin{array}{ccc|ccc|c} 1 & -2 & 1 & 0 & 1 & 0 & -9 \\ 0 & 7 & -4 & 1 & -3 & 0 & 29 \\ 0 & 11 & -2 & 0 & -4 & 1 & 37 \end{array} \right]$$

$$\left[\begin{array}{ccc|ccc|c} 1 & 0 & -\frac{1}{7} & \frac{2}{7} & \frac{1}{7} & 0 & -\frac{5}{7} \\ 0 & 1 & -\frac{4}{7} & \frac{1}{7} & -\frac{3}{7} & 0 & \frac{29}{7} \\ 0 & 0 & \frac{30}{7} & -\frac{11}{7} & \frac{5}{7} & 1 & -\frac{60}{7} \end{array} \right]$$

$$\left[\begin{array}{ccc|ccc|c} 1 & 0 & 0 & \frac{7}{30} & \frac{1}{6} & \frac{1}{30} & -1 \\ 0 & 1 & 0 & -\frac{1}{15} & -\frac{1}{3} & \frac{2}{15} & 3 \\ 0 & 0 & 1 & -\frac{11}{30} & \frac{1}{6} & \frac{7}{30} & -2 \end{array} \right]$$

Thus, reading directly from the tableau,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \\ -2 \end{bmatrix}$$

or, using the inverse matrix, $\mathbf{x} = \mathbf{A}^{-1}\mathbf{c}$,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{7}{30} & \frac{1}{6} & \frac{1}{30} \\ -\frac{1}{15} & -\frac{1}{3} & \frac{2}{15} \\ -\frac{11}{30} & \frac{1}{6} & \frac{7}{30} \end{bmatrix} \begin{bmatrix} 2 \\ -9 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \\ -2 \end{bmatrix}$$

In practice, the solution would be obtained from $\mathbf{x} = \mathbf{A}^{-1}\mathbf{c}$ only if \mathbf{A}^{-1} had been obtained without obtaining the tableau ($\mathbf{I} | \mathbf{A}^{-1} | \mathbf{x}$).

Using Cramer's rule,

$$|\mathbf{A}| = \begin{vmatrix} 3 & 1 & -1 \\ 1 & -2 & 1 \\ 4 & 3 & 2 \end{vmatrix} = -12 - 3 + 4 - (8 + 9 + 2) = -30$$

$$x_1 = \frac{\begin{vmatrix} 2 & 1 & -1 \\ -9 & -2 & 1 \\ 1 & 3 & 2 \end{vmatrix}}{-30} = \frac{-8 + 27 + 1 - (2 + 6 - 18)}{-30} = \frac{30}{-30} = -1$$

$$x_2 = \frac{\begin{vmatrix} 3 & 2 & -1 \\ 1 & -9 & 1 \\ 4 & 1 & 2 \end{vmatrix}}{-30} = \frac{-54 - 1 + 8 - (36 + 3 + 4)}{-30} = \frac{-90}{-30} = 3$$

$$x_3 = \frac{\begin{vmatrix} 3 & 1 & 2 \\ 1 & -2 & -9 \\ 4 & 3 & 1 \end{vmatrix}}{-30} = \frac{-6 + 6 - 36 - (-16 - 81 + 1)}{-30} = \frac{60}{-30} = -2$$

EXAMPLE

Solve the set of simultaneous linear equations

$$x_1 + x_2 - x_3 = 6$$

$$3x_1 - 4x_2 + 2x_3 = -2$$

$$2x_1 + 5x_2 + x_3 = 0$$

Apply the standard procedure to the tableau

$$(A|I|c)$$

to obtain the tableau

$$(I|A^{-1}|x)$$

$$\left[\begin{array}{ccc|ccc|c} 1 & 1 & -1 & 1 & 0 & 0 & 6 \\ 3 & -4 & 2 & 0 & 1 & 0 & -2 \\ 2 & 5 & 1 & 0 & 0 & 1 & 0 \end{array} \right]$$

$$\left[\begin{array}{ccc|ccc|c} 1 & 1 & -1 & 1 & 0 & 0 & 6 \\ 0 & -7 & 5 & -3 & 1 & 0 & -20 \\ 0 & 3 & 3 & -2 & 0 & 1 & -12 \end{array} \right]$$

$$\left[\begin{array}{ccc|ccc|c} 1 & 0 & -\frac{2}{7} & \frac{4}{7} & \frac{1}{7} & 0 & \frac{22}{7} \\ 0 & 1 & -\frac{5}{7} & \frac{3}{7} & -\frac{1}{7} & 0 & \frac{20}{7} \\ 0 & 0 & \frac{36}{7} & -\frac{23}{7} & \frac{3}{7} & 1 & -\frac{144}{7} \end{array} \right]$$

$$\left[\begin{array}{ccc|ccc|c} 1 & 0 & 0 & \frac{7}{18} & \frac{1}{6} & \frac{1}{18} & 2 \\ 0 & 1 & 0 & -\frac{1}{36} & -\frac{1}{12} & \frac{5}{36} & 0 \\ 0 & 0 & 1 & -\frac{23}{36} & \frac{1}{12} & \frac{7}{36} & -4 \end{array} \right]$$

thus, reading directly from the tableau, $x = A^{-1}c$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ -4 \end{bmatrix}$$

or, using the inverse matrix,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{7}{18} & \frac{1}{6} & \frac{1}{18} \\ -\frac{1}{36} & -\frac{1}{12} & \frac{5}{36} \\ -\frac{23}{36} & \frac{1}{12} & \frac{7}{36} \end{bmatrix} \begin{bmatrix} 6 \\ -2 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ -4 \end{bmatrix}$$

Using Cramer's rule,

$$|A| = \begin{vmatrix} 1 & 1 & -1 \\ 3 & -4 & 2 \\ 2 & 5 & 1 \end{vmatrix} = -4 - 15 + 4 - (8 + 10 + 3) = -36$$

$$x_1 = \frac{\begin{vmatrix} 6 & 1 & -1 \\ -2 & -4 & 2 \\ 0 & 5 & 1 \end{vmatrix}}{-36} = \frac{-24 + 10 + 0 - (0 + 60 - 2)}{-36} = \frac{-72}{-36} = 2$$

$$x_2 = \frac{\begin{vmatrix} 1 & 6 & -1 \\ 3 & -2 & 2 \\ 2 & 0 & 1 \end{vmatrix}}{-36} = \frac{-2 + 0 + 24 - (4 + 0 + 18)}{-36} = 0$$

$$x_3 = \frac{\begin{vmatrix} 1 & 1 & 6 \\ 3 & -4 & -2 \\ 2 & 5 & 0 \end{vmatrix}}{-36} = \frac{0 + 90 - 4 - (-48 - 10 + 0)}{-36} = \frac{144}{-36} = -4$$

PROBLEMS

For each of the following sets of simultaneous linear equations, determine whether there is a unique solution and obtain the solution if there is one.

1. $x + 2y = 1$
 $3x + 4y = 2$
2. $x + y - 2z = 5$
 $2x - 4y + 3z = 6$
 $3x - 3y + z = 11$
3. $x_1 + x_2 + x_3 = 3$
 $2x_1 - x_2 - x_3 = 0$
 $3x_1 - 4x_2 - x_3 = 8$
4. $3x - y + 2z = -2$
 $x + y + z = 5$
 $2x - 2y + z = 3$
5. $x - 5y + 6z = 7$
 $3x + 3y - z = 8$
 $2x + 8y - 7z = 1$
6. $x_1 - 3x_2 = -2$
 $2x_1 + 7x_2 = 3$
7. $x_1 - 4x_2 = -1$
 $3x_1 - 2x_2 - x_3 = 0$
 $x_1 + x_3 = 3$
8. $x_1 + x_2 + x_3 = 9$
 $3x_1 + 2x_3 = 17$
 $x_2 + x_3 = 10$
9. $3x - y = 0$
 $2x + 4y = -14$
10. $3x_1 - 4x_2 = -3$
 $6x_1 + x_2 = 3$
11. $4x_1 - 5x_2 - 7x_3 = 15$
 $3x_1 + 2x_2 - 6x_3 = 8$
 $x_1 - 7x_2 - x_3 = 6$
12. $5x_1 - 2x_2 + x_3 = 12$
 $2x_1 + 2x_2 - 3x_3 = 7$
 $x_1 - 6x_2 + 7x_3 = -2$
13. $2x_1 + 3x_2 + x_3 = 0$
 $4x_1 - 8x_2 - 6x_3 = 2$
 $6x_1 + x_2 - x_3 = 0$
14. $x_1 + 5x_2 - 4x_3 = 0$
 $3x_1 - x_2 + 4x_3 = -4$
 $2x_1 + 3x_2 - 8x_3 = 0$
15. $3x_1 - 3x_2 + 4x_3 = -18$
 $4x_1 - 4x_2 + 4x_3 = -24$
 $2x_1 - 2x_2 + 4x_3 = 6$
16. $x_1 + x_2 + 6x_3 = -2$
 $3x_1 + 2x_2 + x_3 = 0$
 $2x_1 + x_2 + 5x_3 = 2$
17. $x_1 + 2x_2 = 0$
 $x_1 + x_2 + x_3 = 2$
 $2x_1 + 2x_2 + 3x_3 = 7$
18. $x_1 + x_2 + x_3 = 4$
 $2x_1 + 3x_2 + 2x_3 = 5$
 $3x_1 + 4x_2 - 3x_3 = -3$

Answers to Odd-Numbered Problems

1. $x = 0, y = \frac{1}{2}$
3. $x_1 = 1, x_2 = 1, x_3 = 1$
5. linearly dependent, no unique solution
7. $x_1 = 1, x_2 = \frac{1}{2}, x_3 = 2$
9. $x = -1, y = -3$
11. inconsistent, no solution
13. $x_1 = -\frac{1}{2}, x_2 = 1, x_3 = -2$
15. inconsistent, no solution
17. $x_1 = -2, x_2 = 1, x_3 = 3$

7.10 Vector Differentiation

In some maximization and minimization problems, expressions involving vectors and matrices must be differentiated. The derivatives of simple expressions involving vectors and matrices can be obtained directly from the definitions. The vector derivatives of linear functions, quadratic forms, and bilinear forms are defined and illustrated in this section. Second- and higher-order vector derivatives are not discussed explicitly, but can be obtained by successive differentiation in the usual manner.

VECTOR DIFFERENTIATION OF A LINEAR FUNCTION

An expression of the form $\mathbf{a}'\mathbf{x}$ where \mathbf{a} is $n \times 1$ and \mathbf{x} is $n \times 1$ is a *linear function*. For the linear function

$$\begin{aligned}\mathbf{a}'\mathbf{x} &= [a_1, a_2, \dots, a_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\ &= a_1x_1 + a_2x_2 + \dots + a_nx_n\end{aligned}$$

the partial derivatives of $\mathbf{a}'\mathbf{x}$ with respect to the scalar x_i , $i = 1, 2, \dots, n$, are

$$\frac{\partial(\mathbf{a}'\mathbf{x})}{\partial x_1} = a_1$$

$$\frac{\partial(\mathbf{a}'\mathbf{x})}{\partial x_2} = a_2$$

$$\vdots$$

$$\frac{\partial(\mathbf{a}'\mathbf{x})}{\partial x_n} = a_n$$

that is, the partial derivatives are the elements of the vector \mathbf{a} . Thus if the n partial derivatives are arranged as a vector \mathbf{a} , the process of vector differentiation is defined by

$$\frac{\partial(\mathbf{a}'\mathbf{x})}{\partial \mathbf{x}} = \mathbf{a}$$

where $\frac{\partial(\mathbf{a}'\mathbf{x})}{\partial \mathbf{x}}$ indicates the operation of differentiating $\mathbf{a}'\mathbf{x}$ with respect to the elements of the vector \mathbf{x} .

EXAMPLE

$$\text{If } \mathbf{a} = \begin{bmatrix} 2a \\ -a \\ 3a \\ a \end{bmatrix} \quad \text{and} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}, \quad \text{then}$$

$$\mathbf{a}'\mathbf{x} = 2ax_1 - ax_2 + 3ax_3 + ax_4$$

and

$$\frac{\partial}{\partial x_1}(\mathbf{a}'\mathbf{x}) = 2a$$

$$\frac{\partial}{\partial x_2}(\mathbf{a}'\mathbf{x}) = -a$$

$$\frac{\partial}{\partial x_3}(\mathbf{a}'\mathbf{x}) = 3a$$

$$\frac{\partial}{\partial x_4}(\mathbf{a}'\mathbf{x}) = a$$

Thus

$$\frac{\partial}{\partial \mathbf{x}}(\mathbf{a}'\mathbf{x}) = \begin{bmatrix} 2a \\ -a \\ 3a \\ a \end{bmatrix} = \mathbf{a}$$

Vector Differentiation of a Vector of Functions

If \mathbf{y} denotes an n -dimensional column vector each element of which is a function of the m elements of \mathbf{x} , that is, if

$$y_i = f_i(x_1, x_2, \dots, x_m) \quad i = 1, 2, \dots, n$$

then each y_i can be differentiated partially with respect to each x_j and these partial derivatives can be arranged in an $m \times n$ matrix as follows:

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} & \dots & \frac{\partial y_n}{\partial x_1} \\ \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} & \dots & \frac{\partial y_n}{\partial x_2} \\ \vdots & \vdots & & \vdots \\ \frac{\partial y_1}{\partial x_m} & \frac{\partial y_2}{\partial x_m} & \dots & \frac{\partial y_n}{\partial x_m} \end{bmatrix}$$

EXAMPLE

If

$$\mathbf{y} = \begin{bmatrix} x_1^2 + 2x_2^2 - 3x_1x_3 \\ 3x_1x_2^2 - x_2^2 + 4x_2x_3^2 \end{bmatrix}$$

then

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \begin{bmatrix} 2x_1 - 3x_3 & 4x_2 & -3x_1 \\ 6x_1x_2 - 2x_2 + 4x_3^2 & 8x_2x_3 & 0 \end{bmatrix}$$

VECTOR DIFFERENTIATION OF A QUADRATIC FORM

An expression of the form $\mathbf{x}'\mathbf{A}\mathbf{x}$, where \mathbf{A} is an $n \times n$ symmetric matrix, is a quadratic form. The quadratic form $\mathbf{x}'\mathbf{A}\mathbf{x}$ can be expanded as follows

$$\begin{aligned} \mathbf{x}'\mathbf{A}\mathbf{x} &= [x_1, x_2, \dots, x_n] \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{12} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\ &= a_{11}x_1^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + \dots + 2a_{1n}x_1x_n \\ &\quad + a_{22}x_2^2 + 2a_{23}x_2x_3 + \dots + 2a_{2n}x_2x_n \\ &\quad + \dots \dots \dots \\ &\quad \quad \quad + a_{nn}x_n^2 \end{aligned}$$

Taking partial derivatives with respect to the elements of \mathbf{x} ,

$$\frac{\partial}{\partial x_1} (\mathbf{x}'\mathbf{A}\mathbf{x}) = 2(a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n)$$

$$\frac{\partial}{\partial x_2} (\mathbf{x}'\mathbf{A}\mathbf{x}) = 2(a_{12}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n)$$

$$\vdots \qquad \qquad \qquad \vdots$$

$$\frac{\partial}{\partial x_n} (\mathbf{x}'\mathbf{A}\mathbf{x}) = 2(a_{1n}x_1 + a_{2n}x_2 + a_{3n}x_3 + \cdots + a_{nn}x_n)$$

Apart from the factor 2, the right-hand side of this set of equations contains the elements of $\mathbf{A}\mathbf{x}$, which is an n -dimensional column vector; alternatively, the right-hand side contains the elements of $\mathbf{x}'\mathbf{A}$, which is an n -dimensional row vector. Thus

$$\frac{\partial}{\partial \mathbf{x}} (\mathbf{x}'\mathbf{A}\mathbf{x}) = 2\mathbf{A}\mathbf{x}$$

or

$$\frac{\partial}{\partial \mathbf{x}} (\mathbf{x}'\mathbf{A}\mathbf{x}) = 2\mathbf{x}'\mathbf{A}$$

In practice, the choice between these two forms usually depends on the context in which differentiation occurs, since matrices can be equated only if they are the same order and, specifically, a row vector cannot be set equal to a column vector.

EXAMPLE

$$\text{If } \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \text{and} \quad \mathbf{A} = \begin{bmatrix} 3 & 1 & -2 \\ 1 & 0 & 3 \\ -2 & 3 & 2 \end{bmatrix}, \quad \text{then}$$

$$\begin{aligned} \mathbf{x}'\mathbf{A}\mathbf{x} &= [x_1, x_2, x_3] \begin{bmatrix} 3 & 1 & -2 \\ 1 & 0 & 3 \\ -2 & 3 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\ &= [3x_1 + x_2 - 2x_3, \quad x_1 + 3x_3, \quad -2x_1 + 3x_2 + 2x_3] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \end{aligned}$$

$$= 3x_1^2 + x_1x_2 - 2x_1x_3 + x_1x_2 + 3x_2x_3 - 2x_1x_3 + 3x_2x_3 + 2x_3^2$$

$$= 3x_1^2 + 2x_3^2 + 2x_1x_2 - 4x_1x_3 + 6x_2x_3$$

and

$$\frac{\partial}{\partial x_1} (\mathbf{x}'\mathbf{A}\mathbf{x}) = 6x_1 + 2x_2 - 4x_3$$

$$\frac{\partial}{\partial x_2} (\mathbf{x}'\mathbf{A}\mathbf{x}) = 2x_1 + 6x_3$$

$$\frac{\partial}{\partial x_3} (\mathbf{x}'\mathbf{A}\mathbf{x}) = 4x_3 - 4x_1 + 6x_2$$

Thus

$$\frac{\partial}{\partial \mathbf{x}} (\mathbf{x}' \mathbf{A} \mathbf{x}) = 2 \begin{bmatrix} 3x_1 + x_2 - 2x_3 \\ x_1 + 3x_3 \\ -2x_1 + 3x_2 + 2x_3 \end{bmatrix} = 2\mathbf{A}\mathbf{x}$$

or

$$\frac{\partial}{\partial \mathbf{x}} (\mathbf{x}' \mathbf{A} \mathbf{x}) = 2[3x_1 + x_2 - 2x_3, \quad x_1 + 3x_3, \quad -2x_1 + 3x_2 + 2x_3] = 2\mathbf{x}' \mathbf{A}$$

VECTOR DIFFERENTIATION OF A BILINEAR FORM

An expression of the form $\mathbf{x}' \mathbf{B} \mathbf{z}$, where \mathbf{x}' is $1 \times m$, \mathbf{B} is $m \times n$ and \mathbf{z} is $n \times 1$, is a bilinear form. The following derivatives for a bilinear form can be verified using the procedures above.

$$\frac{\partial}{\partial \mathbf{x}} (\mathbf{x}' \mathbf{B} \mathbf{z}) = \mathbf{B} \mathbf{z}$$

$$\frac{\partial}{\partial \mathbf{z}} (\mathbf{x}' \mathbf{B} \mathbf{z}) = \mathbf{B}' \mathbf{x}$$

EXAMPLE

If

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 2b_1 & b_1 + b_2 \\ b_2 + 2b_3 & 3b_2 \\ b_1 + b_3 & 4b_3^2 \end{bmatrix} \quad \mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

then

$$\begin{aligned} \mathbf{x}' \mathbf{B} \mathbf{z} &= [x_1, x_2, x_3] \begin{bmatrix} 2b_1 & b_1 + b_2 \\ b_2 + 2b_3 & 3b_2 \\ b_1 + b_3 & 4b_3^2 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \\ &= [2b_1 x_1 + b_2 x_2 + 2b_3 x_2 + b_1 x_3 + b_3 x_3, \quad b_1 x_1 + b_2 x_1 + 3b_2 x_2 + 4b_3^2 x_3] \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \\ &= 2b_1 x_1 z_1 + b_2 x_2 z_1 + 2b_3 x_2 z_1 + b_1 x_3 z_1 + b_3 x_3 z_1 + b_1 x_1 z_2 \\ &\quad + b_2 x_1 z_2 + 3b_2 x_2 z_2 + 4b_3^2 x_3 z_2 \end{aligned}$$

and

$$\frac{\partial}{\partial x_1} (\mathbf{x}' \mathbf{B} \mathbf{z}) = 2b_1 z_1 + b_1 z_2 + b_2 z_2$$

$$\frac{\partial}{\partial x_2} (\mathbf{x}' \mathbf{B} \mathbf{z}) = b_2 z_1 + 2b_3 z_1 + 3b_2 z_2$$

$$\frac{\partial}{\partial x_3} (\mathbf{x}' \mathbf{B} \mathbf{z}) = b_1 z_1 + b_3 z_1 + 4b_3^2 z_2$$

Thus

$$\frac{\partial}{\partial \mathbf{x}} (\mathbf{x}' \mathbf{B} \mathbf{z}) = \begin{bmatrix} 2b_1 z_1 + b_1 z_2 + b_2 z_2 \\ b_2 z_1 + 2b_3 z_1 + 3b_2 z_2 \\ b_1 z_1 + b_3 z_1 + 4b_3^2 z_2 \end{bmatrix} = \mathbf{B} \mathbf{z}$$

and

$$\frac{\partial}{\partial z_1} (\mathbf{x}'\mathbf{B}\mathbf{z}) = 2b_1x_1 + b_2x_2 + 2b_3x_2 + b_1x_3 + b_3x_3$$

$$\frac{\partial}{\partial z_2} (\mathbf{x}'\mathbf{B}\mathbf{z}) = b_1x_1 + b_2x_1 + 3b_2x_2 + 4b_3^2x_3$$

Thus

$$\begin{aligned} \frac{\partial}{\partial \mathbf{z}} (\mathbf{x}'\mathbf{B}\mathbf{z}) &= \begin{bmatrix} 2b_1x_1 + b_2x_2 + 2b_3x_2 + b_1x_3 + b_3x_3 \\ b_1x_1 + b_2x_1 + 3b_2x_2 + 4b_3^2x_3 \end{bmatrix} \\ &= \begin{bmatrix} 2b_1 & b_2 + 2b_3 & b_1 + b_3 \\ b_1 + b_2 & 3b_2 & 4b_3^2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\ &= \mathbf{B}'\mathbf{x} \end{aligned}$$

APPLICATIONS OF VECTOR DIFFERENTIATION IN MAXIMIZATION AND MINIMIZATION

Vector differentiation is used in many areas of applications for problems of maximization and minimization. In particular, the derivations of most of the estimators in multivariate analysis involve vector differentiation. As an example of these applications, the derivation of multivariate regressions estimators is discussed in this section.

In multivariate regression analysis, the model involves the assumption of a linear relationship between a variable y and k explanatory variables x_1, x_2, \dots, x_k and a disturbance term u . For a sample of n observations on y and the x 's,

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + u_i \quad i = 1, 2, \dots, n$$

and the n equations can be written in matrix notation as

$$\mathbf{y} = \mathbf{x}\boldsymbol{\beta} + \mathbf{u}$$

where

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} 1 & x_{21} & \cdots & x_{k1} \\ 1 & x_{22} & \cdots & x_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{2n} & \cdots & x_{kn} \end{bmatrix}$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} \quad \mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

The intercept β_0 is represented by the column of ones in the \mathbf{x} matrix. As a result of the convention of denoting the i th observation on the k th variable by x_{ki} , the subscripts in the \mathbf{x} matrix are the reverse of the usual pattern, which indicates the row by the first subscript and the column by the second subscript.

The problem is to obtain estimates of the β 's. These estimates are customarily obtained using the least squares principle. Let

$$\hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_k \end{bmatrix}$$

denote a vector of estimates of β . Then $y = x\hat{\beta} + e$ where e denotes the column vector of n residuals $(y - x\hat{\beta})$. The least squares principle minimizes the sum of squared residuals

$$\begin{aligned}\sum_{i=1}^n e_i^2 &= e'e \\ &= (y - x\hat{\beta})'(y - x\hat{\beta}) \\ &= y'y - 2\hat{\beta}'x'y + \hat{\beta}'x'x\hat{\beta}\end{aligned}$$

Note that $(y - x\hat{\beta})' = y' - \hat{\beta}'x'$ and $\hat{\beta}'x'y$ is a scalar and thus equal to its transpose $y'x\hat{\beta}$. To obtain the estimate $\hat{\beta}$ which minimizes the sum of squared residuals, differentiate $e'e$ with respect to $\hat{\beta}$,

$$\begin{aligned}\frac{\partial}{\partial \hat{\beta}}(e'e) &= -2x'y + x'x\hat{\beta} + (\hat{\beta}'x'x)' \\ &= -2x'y + 2x'x\hat{\beta}\end{aligned}$$

$$\text{If } \frac{\partial}{\partial \hat{\beta}}(e'e) = 0.$$

$$x'x\hat{\beta} = x'y$$

and

$$\hat{\beta} = (x'x)^{-1}x'y$$

Alternatively,

$$\frac{\partial}{\partial \hat{\beta}}(e'e) = -2y'x + 2\hat{\beta}'x'x$$

$$\text{If } \frac{\partial}{\partial \hat{\beta}}(e'e) = 0,$$

$$\hat{\beta}'x'x = y'x$$

$$\hat{\beta}' = y'x[x'x]^{-1}$$

and

$$\hat{\beta} = [x'x]^{-1}x'y$$

is the vector of least squares estimates for the regression coefficients.

NOTE: To show that $\hat{\beta} = [x'x]^{-1}x'y$ is a minimizing solution, the second derivative must be examined

$$\begin{aligned}\frac{\partial^2}{\partial \hat{\beta}^2}[e'e] &= \frac{\partial}{\partial \hat{\beta}}[-2x'y + 2x'x\hat{\beta}] \\ &= 2x'x\end{aligned}$$

Thus the second order condition for minimization requires that $x'x$ is positive definite, that is, that $y'[x'x]y > 0$ for all $y > 0$. In practical examples, $x'x$ is positive definite.

EXAMPLE

$$\text{If } \mathbf{x} = \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & -2 \end{bmatrix} \quad \mathbf{y} = \begin{bmatrix} 5 \\ 3 \\ -2 \\ 8 \end{bmatrix}$$

find the least squares estimate $\hat{\beta}$ for the regression $\mathbf{y} = \mathbf{x}\beta + \mathbf{u}$.

$$\mathbf{x}'\mathbf{x} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & -2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & -2 \end{bmatrix}$$

$$= \begin{bmatrix} 4 & -2 \\ -2 & 6 \end{bmatrix}$$

$$|\mathbf{x}'\mathbf{x}| = 24 - 4 = 20$$

$$[\mathbf{x}'\mathbf{x}]^{-1} = \begin{bmatrix} \frac{3}{10} & \frac{1}{10} \\ \frac{1}{10} & \frac{1}{5} \end{bmatrix}$$

and

$$\begin{aligned} \hat{\beta} &= [\mathbf{x}'\mathbf{x}]^{-1}\mathbf{x}'\mathbf{y} = \begin{bmatrix} \frac{3}{10} & \frac{1}{10} \\ \frac{1}{10} & \frac{1}{5} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & -2 \end{bmatrix} \begin{bmatrix} 5 \\ 3 \\ -2 \\ 8 \end{bmatrix} \\ &= \begin{bmatrix} \frac{19}{10} \\ -\frac{16}{5} \end{bmatrix} \end{aligned}$$

Thus the estimated regression is $\hat{y} = \frac{19}{10} - \frac{16}{5}x$.

PROBLEMS

1. If $\mathbf{b} = \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, find $\frac{\partial}{\partial \mathbf{x}}[\mathbf{b}'\mathbf{x}]$.
2. If $\mathbf{c} = \begin{bmatrix} -c \\ 0 \\ 2c \\ -3c \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}$, find $\frac{\partial}{\partial \mathbf{y}}[\mathbf{c}'\mathbf{y}]$.
3. If $\mathbf{c} = \begin{bmatrix} c^2 + 1 \\ 3c \\ 4c - 5 \\ c^3 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$, find $\frac{\partial}{\partial \mathbf{x}}(\mathbf{c}'\mathbf{x})$.
4. If $\mathbf{y} = \begin{bmatrix} 2x_1^2 - 3x_2x_3 \\ x_1 - 4x_1x_2x_3 \\ x_1^2 - 2x_2 + x_3 \end{bmatrix}$, find $\frac{\partial \mathbf{y}}{\partial \mathbf{x}}$.
5. If $\mathbf{y} = \begin{bmatrix} x_1^2 + 3x_2 \\ 2x_1x_2 - x_2 \\ 2x_1 + x_1x_2 - 3x_2^2 \\ 3x_1^2 - x_1^2x_2 - x_2^3 \end{bmatrix}$, find $\frac{\partial \mathbf{y}}{\partial \mathbf{x}}$.

6. If $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 3 + b_1 & b_1 + 2b_2 & 4b_2^2 \\ b_2^2 & 4b_1 & b_1 b_2 \end{bmatrix}$

and $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$, find $\frac{\partial}{\partial \mathbf{x}} [\mathbf{x}' \mathbf{B} \mathbf{y}]$ and $\frac{\partial}{\partial \mathbf{y}} [\mathbf{x}' \mathbf{B} \mathbf{y}]$.

7. If $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, $\mathbf{A} = \begin{bmatrix} 3 & 0 & -1 \\ 2 & 1 & 4 \\ -1 & 0 & 3 \end{bmatrix}$

and $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$, find $\frac{\partial}{\partial \mathbf{x}} [\mathbf{x}' \mathbf{A} \mathbf{y}]$ and $\frac{\partial}{\partial \mathbf{y}} [\mathbf{x}' \mathbf{A} \mathbf{y}]$.

8. If $\mathbf{x} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} -3 \\ 5 \\ 11 \end{bmatrix}$,

find the least squares estimate $\hat{\boldsymbol{\beta}}$ for the regression $\mathbf{y} = \mathbf{x}\boldsymbol{\beta} + \mathbf{u}$.

Answers to Odd-Numbered Problems

1. $\begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}$

3. $\begin{bmatrix} c^2 + 1 \\ 3c \\ 4c - 5 \\ c^3 \end{bmatrix}$

5. $\begin{bmatrix} 2x_1 & 2x_2 & 2 + x_2 & 6x_1 - 2x_1 x_2 \\ 3 & 2x_1 - 1 & x_1 - 6x_2 & -x_1 - 3x_2^2 \end{bmatrix}$

7. $\begin{bmatrix} 3y_1 - y_3 \\ 2y_1 + y_2 + 4y_3 \\ -y_1 + 3y_3 \end{bmatrix}$

Applications of Matrix Algebra

8

8.1 Introduction

There are many applications of matrix algebra in various areas. The applications discussed in this chapter are maxima and minima of functions of n variables, input-output analysis, linear programming, game theory, and Markov processes.

8.2 Maxima and Minima of Functions of n Variables

The necessary and sufficient conditions for a point to be a maximum or minimum of a function of two variables are discussed in previous sections. These results are a special case ($n = 2$) of the conditions for a maximum or minimum of a function of n variables; the general results are given below in matrix notation. Unconstrained maximization and minimization and constrained maximization and minimization using Lagrange multipliers and the Kuhn-Tucker conditions are discussed and illustrated.

UNCONSTRAINED MAXIMA AND MINIMA

Consider a function of n variables $f(x_1, x_2, \dots, x_n)$ at a point $x^* = (x_1^*, x_2^*, \dots, x_n^*)$ such that the n partial derivatives are zero:

$$\left. \frac{\partial f}{\partial x_1} \right|_{x^*} = 0 \quad \left. \frac{\partial f}{\partial x_2} \right|_{x^*} = 0, \quad \dots, \quad \left. \frac{\partial f}{\partial x_n} \right|_{x^*} = 0$$

Define the determinant Δ_n of second-order partial derivatives

$$\Delta_n = \begin{vmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \dots & \frac{\partial^2 f}{\partial x_n^2} \end{vmatrix}$$

sometimes referred to as a *Hessian determinant*.

The principal minors of Δ_n are

$$\Delta_1 = \frac{\partial^2 f}{\partial x_1^2}$$

$$\Delta_2 = \begin{vmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{vmatrix}$$

$$\Delta_3 = \begin{vmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_1 \partial x_3} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \frac{\partial^2 f}{\partial x_2 \partial x_3} \\ \frac{\partial^2 f}{\partial x_3 \partial x_1} & \frac{\partial^2 f}{\partial x_3 \partial x_2} & \frac{\partial^2 f}{\partial x_3^2} \end{vmatrix}$$

\vdots

$$\Delta_n =$$

The stationary point $x^* = (x_1^*, x_2^*, \dots, x_n^*)$ is a

Local maximum if $\Delta_1 < 0, \Delta_2 > 0, \Delta_3 < 0, \dots$

Local minimum if $\Delta_1 > 0, \Delta_2 > 0, \Delta_3 > 0, \dots$

If neither of these conditions holds, the function must be examined in the region of the stationary point. Note that the conditions given previously for maxima and minima of a function of two variables are a special case of those given above for n variables.

EXAMPLE

Determine the maxima or minima (if any) of the function

$$f(x_1, x_2, x_3) = x_1^2 + 2x_2^2 + x_3^2 + x_1 x_2 - 2x_3 - 7x_1 + 12\Delta_n$$

$$\frac{\partial f}{\partial x_1} = 2x_1 + x_2 - 7$$

$$\frac{\partial f}{\partial x_2} = 4x_2 + x_1$$

$$\frac{\partial f}{\partial x_3} = 2x_3 - 2$$

If $\frac{\partial f}{\partial x_1} = 0, \frac{\partial f}{\partial x_2} = 0, \frac{\partial f}{\partial x_3} = 0$, then

$$2x_1 + x_2 = 7$$

$$x_1 + 4x_2 = 0$$

$$2x_3 = 2$$

Thus

$$x_2 = -1$$

$$x_1 = 4$$

$$x_3 = 1$$

$$\frac{\partial^2 f}{\partial x_1^2} = 2 \quad \frac{\partial^2 f}{\partial x_1 \partial x_2} = 1 \quad \frac{\partial^2 f}{\partial x_1 \partial x_3} = 0$$

$$\frac{\partial^2 f}{\partial x_2^2} = 4 \quad \frac{\partial^2 f}{\partial x_2 \partial x_3} = 0$$

$$\frac{\partial^2 f}{\partial x_3^2} = 2$$

$$\Delta_1 = 2$$

$$\Delta_2 = \begin{vmatrix} 2 & 1 \\ 1 & 4 \end{vmatrix} = 8 - 1 = 7$$

$$\Delta_3 = \begin{vmatrix} 2 & 1 & 0 \\ 1 & 4 & 0 \\ 0 & 0 & 2 \end{vmatrix} = 16 + 0 + 0 - (0 + 0 + 2) = 14$$

$\Delta_1 > 0$, $\Delta_2 > 0$, $\Delta_3 > 0$, so the point $(4, -1, 1)$ is a local minimum of the function $f(x_1, x_2, x_3) = x_1^2 + 2x_2^2 + x_3^2 + x_1 x_2 - 2x_3 - 7x_1 + 12$.

EXAMPLE

Determine the maxima and minima (if any) of the function

$$f(x, y, z) = e^{-x^2 - y^2 - z^2 + 2y + xz}$$

$$\frac{\partial f}{\partial x} = (-2x + z)e^{-x^2 - y^2 - z^2 + 2y + xz}$$

$$\frac{\partial f}{\partial y} = (-2y + 2)e^{-x^2 - y^2 - z^2 + 2y + xz}$$

$$\frac{\partial f}{\partial z} = (-2z + x)e^{-x^2 - y^2 - z^2 + 2y + xz}$$

If $\frac{\partial f}{\partial x} = 0$, $\frac{\partial f}{\partial y} = 0$, $\frac{\partial f}{\partial z} = 0$, then

$$-2x + z = 0$$

$$2y = 2$$

$$x - 2z = 0$$

Thus

$$x = z = 0$$

$$y = 1$$

$$\frac{\partial^2 f}{\partial x^2} = [(-2x + z)^2 - 2]e^{-x^2 - y^2 - z^2 + 2y + xz}$$

$$\left. \frac{\partial^2 f}{\partial x^2} \right|_{(0, 1, 0)} = -2e$$

$$\frac{\partial^2 f}{\partial x \partial y} = (-2x + z)(-2y + 2)e^{-x^2 - y^2 - z^2 + 2y + xz}$$

$$\left. \frac{\partial^2 f}{\partial x \partial y} \right|_{(0, 1, 0)} = 0$$

$$\frac{\partial^2 f}{\partial x \partial z} = [(2x + z)(-2z + x) + 1]e^{-x^2 - y^2 - z^2 + 2y + xz}$$

$$\left. \frac{\partial^2 f}{\partial x \partial z} \right|_{(0, 1, 0)} = e$$

$$\frac{\partial^2 f}{\partial y^2} = [(-2y + 2)^2 - 2]e^{-x^2 - y^2 - z^2 + 2y + xz}$$

$$\left. \frac{\partial^2 f}{\partial y^2} \right|_{(0, 1, 0)} = -2e$$

$$\frac{\partial^2 f}{\partial y \partial z} = (-2y + 2)(-2z + x)e^{-x^2 - y^2 - z^2 + 2y + xz}$$

$$\left. \frac{\partial^2 f}{\partial y \partial z} \right|_{(0, 1, 0)} = 0$$

$$\frac{\partial^2 f}{\partial z^2} = [(-2z + x)^2 - 2]e^{-x^2 - y^2 - z^2 + 2y + xz}$$

$$\left. \frac{\partial^2 f}{\partial z^2} \right|_{(0, 1, 0)} = -2e$$

$$\Delta_1 = -2e$$

$$\Delta_2 = \begin{vmatrix} -2e & 0 \\ 0 & -2e \end{vmatrix} = 4e$$

$$\Delta_3 = \begin{vmatrix} -2e & 0 & e \\ 0 & -2e & 0 \\ e & 0 & -2e \end{vmatrix} = -8e^3 + 2e^3 = -6e^3$$

$\Delta_1 < 0$, $\Delta_2 > 0$, $\Delta_3 < 0$, so the point $(0, 1, 0)$ is a local maximum of the function $f(x, y, z) = e^{-x^2 - y^2 - z^2 + 2y + xz}$.

EXAMPLE

Total profit is equal to total revenue minus total cost. In a perfectly competitive market, total revenue is equal to the product of the number of units sold and the fixed unit price; thus

$$P = pq - C$$

where P is profit, p is price per unit, q is quantity sold, and C is total cost. If the production function is

$$q = f(x_1, x_2)$$

where x_1 and x_2 are the variable inputs, and the cost function is

$$C = p_1 x_1 + p_2 x_2 + b$$

where p_1 and p_2 are the unit prices of x_1 and x_2 , respectively, and b is the cost of the fixed inputs, then

$$P = pf(x_1, x_2) - p_1 x_1 - p_2 x_2 - b$$

Profit is thus a function of x_1 and x_2 and is maximized with respect to these variables as follows.

$$\frac{\partial P}{\partial x_1} = pf_1 - p_1 = 0$$

$$\frac{\partial P}{\partial x_2} = pf_2 - p_2 = 0$$

or

$$pf_1 = p_1$$

$$pf_2 = p_2$$

Since f_1 and f_2 are the marginal products of the two inputs, pf_1 and pf_2 are the values of these marginal products, that is, the rates at which revenue would increase with increase of the respective inputs. Thus the first-order conditions for profit maximization require that each input be increased until the value of its marginal product equals its price. Note that the

maximum profit-input combination lies on the expansion path since the conditions $pf_1 = p_1$ and $pf_2 = p_2$ are a special case of the condition $f_1/f_2 = p_1/p_2$.

Second-order conditions require that the principal minors of the Hessian determinant

$$\begin{vmatrix} \frac{\partial^2 P}{\partial x_1^2} & \frac{\partial^2 P}{\partial x_1 \partial x_2} \\ \frac{\partial^2 P}{\partial x_2 \partial x_1} & \frac{\partial^2 P}{\partial x_2^2} \end{vmatrix}$$

alternate in sign. That is, for a maximum it is required that

$$\begin{aligned} \frac{\partial^2 P}{\partial x_1^2} &< 0 \\ \begin{vmatrix} \frac{\partial^2 P}{\partial x_1^2} & \frac{\partial^2 P}{\partial x_1 \partial x_2} \\ \frac{\partial^2 P}{\partial x_2 \partial x_1} & \frac{\partial^2 P}{\partial x_2^2} \end{vmatrix} &> 0 \end{aligned}$$

Expanding the determinant, the second condition is

$$\left(\frac{\partial^2 P}{\partial x_1^2} \right) \left(\frac{\partial^2 P}{\partial x_2^2} \right) - \left(\frac{\partial^2 P}{\partial x_1 \partial x_2} \right)^2 > 0$$

Since $\frac{\partial^2 P}{\partial x_1^2} < 0$ (first condition) and $\left(\frac{\partial^2 P}{\partial x_1 \partial x_2} \right)^2 > 0$,

$$\frac{\partial^2 P}{\partial x_2^2} < 0$$

The second-order partials can be evaluated as follows:

$$\frac{\partial^2 P}{\partial x_1^2} = pf_{11} < 0$$

$$\frac{\partial^2 P}{\partial x_2^2} = pf_{22} < 0$$

and, since $p > 0$, the second-order conditions for maximization require that

$$f_{11} < 0$$

$$f_{22} < 0$$

Thus the second-order conditions for maximization require that both marginal products be decreasing. This is reasonable, since the price of a variable is constant under free competition and thus, if the marginal product of a variable were increasing, profit could be increased by increasing the quantity of the variable.

PROBLEMS

1. Determine the values of x_1 , x_2 , and x_3 (if any) that maximize or minimize the function

$$f(x_1, x_2, x_3) = x_1 x_2 + 10x_1 - x_1^2 - x_2^2 - x_3^2.$$

2. Determine the values of x , y , and z (if any) that maximize or minimize the function

$$f(x, y, z) = e^{4x^2 + 2y^2 + z^2 - 5xy - 4z}.$$

3. Determine the values of x , y , and z (if any) that maximize or minimize the function

$$f(x, y, z) = e^{-x^2 - 2y^2 - z^2 - 2xy}.$$

4. Determine the values of x_1 , x_2 , and x_3 (if any) that maximize or minimize the function

$$f(x_1, x_2, x_3) = x_1^2 + x_2^2 + 7x_3^2 - x_1x_2.$$

Answers to Odd-Numbered Problems

1. $x_1 = \frac{20}{3}$, $x_2 = \frac{10}{3}$, $x_3 = 0$ (maximize)
 3. $x = y = z = 0$ (maximize)

LAGRANGE MULTIPLIERS

Consider a function of n variables $f(x_1, x_2, \dots, x_n)$ subject to the constraint $g(x_1, x_2, \dots, x_n) = 0$ at a point $x^* = (x_1^*, x_2^*, \dots, x_n^*)$ which satisfies the $n + 1$ equations

$$\frac{\partial f}{\partial x_1} - \lambda \frac{\partial g}{\partial x_1} = 0$$

$$\frac{\partial f}{\partial x_2} - \lambda \frac{\partial g}{\partial x_2} = 0$$

$$\vdots$$

$$\frac{\partial f}{\partial x_n} - \lambda \frac{\partial g}{\partial x_n} = 0$$

$$\frac{\partial f}{\partial \lambda} - \lambda \frac{\partial g}{\partial \lambda} = 0$$

Note that for $x^* = (x_1^*, x_2^*, \dots, x_n^*)$, $\lambda = (f_{x_i}/g_{x_i})$, $i = 1, 2, \dots, n$. Define the determinant

$$\Delta_{n+1} = \begin{vmatrix} 0 & g_{x_1} & g_{x_2} & \cdots & g_{x_n} \\ g_{x_1} & f_{x_1x_1} - \lambda g_{x_1x_1} & f_{x_1x_2} - \lambda g_{x_1x_2} & \cdots & f_{x_1x_n} - \lambda g_{x_1x_n} \\ g_{x_2} & f_{x_2x_1} - \lambda g_{x_2x_1} & f_{x_2x_2} - \lambda g_{x_2x_2} & \cdots & f_{x_2x_n} - \lambda g_{x_2x_n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ g_{x_n} & f_{x_nx_1} - \lambda g_{x_nx_1} & f_{x_nx_2} - \lambda g_{x_nx_2} & \cdots & f_{x_nx_n} - \lambda g_{x_nx_n} \end{vmatrix}$$

sometimes referred to as a *bordered Hessian determinant*. In order to determine whether $x^* = (x_1^*, x_2^*, \dots, x_n^*)$ is a maximum or minimum, the $n - 1$ principal minors of Δ_{n+1} : $\Delta_3, \Delta_4, \dots, \Delta_{n+1}$, must be evaluated for $x^* = (x_1^*, x_2^*, \dots, x_n^*)$. Note that Δ_i consists of the first i rows and i columns of Δ_{n+1} . The point $x^* = (x_1^*, x_2^*, \dots, x_n^*)$ is a

Local (constrained) maximum if $\Delta_3 > 0$, $\Delta_4 < 0$, $\Delta_5 > 0$, ...

Local (constrained) minimum if $\Delta_3 < 0$, $\Delta_4 < 0$, $\Delta_5 < 0$, ...

If neither of these conditions holds, the function must be examined in the region of the stationary point. Note that the conditions given previously for constrained maxima and minima of a function of two variables are a special case of those given above for n variables.

EXAMPLE

Determine the minimum of the function

$$f(x_1, x_2, x_3) = x_1 x_2 + x_1 x_3 + x_2 x_3 \quad \text{if } x_1 x_2 x_3 = 125.$$

$$F(x_1, x_2, x_3, \lambda) = x_1 x_2 + x_1 x_3 + x_2 x_3 - \lambda(x_1 x_2 x_3 - 125)$$

$$\frac{\partial F}{\partial x_1} = x_2 + x_3 - \lambda x_2 x_3$$

$$\frac{\partial F}{\partial x_2} = x_1 + x_3 - \lambda x_1 x_3$$

$$\frac{\partial F}{\partial x_3} = x_1 + x_2 - \lambda x_1 x_2$$

$$\frac{\partial F}{\partial \lambda} = -(x_1 x_2 x_3 - 125)$$

If $\frac{\partial F}{\partial x_1} = 0$, $\frac{\partial F}{\partial x_2} = 0$, $\frac{\partial F}{\partial x_3} = 0$, $\frac{\partial F}{\partial \lambda} = 0$, then

$$x_1 = x_2 = x_3 = \sqrt[3]{125} = 5 \quad \lambda = \frac{2}{5}$$

$$f_{x_1 x_1} = 0 \quad f_{x_2 x_2} = 0 \quad f_{x_3 x_3} = 0$$

$$f_{x_1 x_2} = 1 \quad f_{x_1 x_3} = 1 \quad f_{x_2 x_3} = 1$$

$$g_{x_1} = x_2 x_3 = 25 \quad g_{x_2} = x_1 x_3 = 25 \quad g_{x_3} = x_1 x_2 = 25$$

$$g_{x_1 x_1} = 0 \quad g_{x_2 x_2} = 0 \quad g_{x_3 x_3} = 0$$

$$g_{x_1 x_2} = x_3 = 5 \quad g_{x_1 x_3} = x_2 = 5 \quad g_{x_2 x_3} = x_1 = 5$$

$$\Delta_3 = \begin{vmatrix} 0 & 25 & 25 \\ 25 & 0 & -1 \\ 25 & -1 & 0 \end{vmatrix} = 0 - 625 - 625 - (0 + 0 + 0) = -1250$$

$$\Delta_4 = \begin{vmatrix} 0 & 25 & 25 & 25 \\ 25 & 0 & -1 & -1 \\ 25 & -1 & 0 & -1 \\ 25 & -1 & -1 & 0 \end{vmatrix}$$

$$= 0 - 25 \begin{vmatrix} 25 & -1 & -1 \\ 25 & 0 & -1 \\ 25 & -1 & 0 \end{vmatrix} + 25 \begin{vmatrix} 25 & 0 & -1 \\ 25 & -1 & -1 \\ 25 & -1 & 0 \end{vmatrix} - 25 \begin{vmatrix} 25 & 0 & -1 \\ 25 & -1 & 0 \\ 25 & -1 & -1 \end{vmatrix}$$

$$= 0 - 25[0 + 25 + 25 - (0 + 25 + 0)] + 25[0 + 25 + 0 - (25 + 25 + 0)]$$

$$- 25[25 + 25 + 0 - (25 + 0 + 0)]$$

$$= 0 - 625 - 625 - 625 = -1875$$

$\Delta_3 < 0$, $\Delta_4 < 0$, so the point $(5, 5, 5)$ is a local minimum of the function $f(x_1, x_2, x_3) = x_1 x_2 + x_1 x_3 + x_2 x_3$ subject to $x_1 x_2 x_3 = 125$.

EXAMPLE

Determine the maximum of the function

$$f(x, y, z) = -x^2 - 2y^2 - z^2 + xy + z \quad \text{if } x + y + z = 35.$$

$$F(x, y, z, \lambda) = -x^2 - 2y^2 - z^2 + xy + z - \lambda(x + y + z - 35)$$

$$\frac{\partial F}{\partial x} = -2x + y - \lambda$$

$$\frac{\partial F}{\partial y} = -4y + x - \lambda$$

$$\frac{\partial F}{\partial z} = -2z + 1 - \lambda$$

$$\frac{\partial F}{\partial \lambda} = -(x + y + z - 35)$$

If $\frac{\partial F}{\partial x} = 0$, $\frac{\partial F}{\partial y} = 0$, $\frac{\partial F}{\partial z} = 0$, $\frac{\partial F}{\partial \lambda} = 0$, then

$$-2x + y = \lambda$$

$$-4y + x = \lambda$$

$$-2z + 1 = \lambda$$

From the first and second equations,

$$3x = 5y$$

$$x = \frac{5}{3}y$$

From the second and third equations,

$$-2z + 1 = y - \frac{10}{3}y = -\frac{7}{3}y$$

$$z = \frac{1 + \frac{7}{3}y}{2}$$

$$\frac{5}{3}y + y + \frac{1 + \frac{7}{3}y}{2} = 35$$

$$\frac{10}{3}y + 2y + 1 + \frac{7}{3}y = 70$$

$$\frac{23}{3}y = 69$$

$$y = 9$$

$$x = 15$$

$$z = 11$$

$$\lambda = -21$$

$$f_{xx} = -2 \quad f_{yy} = -4 \quad f_{zz} = -2$$

$$f_{xy} = 1 \quad f_{xz} = 0 \quad f_{yz} = 0$$

$$g_x = 1 \quad g_y = 1 \quad g_z = 1$$

$$g_{xx} = 0 \quad g_{yy} = 0 \quad g_{zz} = 0$$

$$g_{xy} = 0 \quad g_{xz} = 0 \quad g_{yz} = 0$$

$$\Delta_3 = \begin{vmatrix} 0 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -4 \end{vmatrix} = (0 + 1 + 1) - (-2 + 0 - 4) = 8$$

$$\begin{aligned}
\Delta_4 &= \begin{vmatrix} 0 & 1 & 1 & 1 \\ 1 & -2 & 1 & 1 \\ 1 & 1 & -4 & 0 \\ 1 & 1 & 0 & -2 \end{vmatrix} \\
&= 0 - \begin{vmatrix} 1 & 1 & 1 \\ 1 & -4 & 0 \\ 1 & 0 & -2 \end{vmatrix} + \begin{vmatrix} 1 & -2 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & -2 \end{vmatrix} - \begin{vmatrix} 1 & -2 & 1 \\ 1 & 1 & -4 \\ 1 & 1 & 0 \end{vmatrix} \\
&= 0 - [8 + 0 + 0 - (-4 + 0 - 2)] + [-2 + 1 + 0 - (1 + 0 + 4)] \\
&\quad - [0 + 1 + 8 - (1 - 4 + 0)] \\
&= 0 - 14 - 6 - 11 = -31
\end{aligned}$$

$\Delta_3 > 0$, $\Delta_4 < 0$, so the point (15, 9, 11) is a local maximum of the function $f(x, y, z) = -x^2 - 2y^2 - z^2 + xy + z$ subject to $x + y + z = 35$.

EXAMPLE

If utility analysis is generalized to the case of n commodities, the utility function is given by

$$U = f(q_1, q_2, \dots, q_n)$$

where q_1, q_2, \dots, q_n are the quantities of the n commodities, and the budget constraint is given by

$$y - \sum_{i=1}^n p_i q_i = 0$$

where p_1, p_2, \dots, p_n are the unit prices of the n commodities. Then the function to be maximized is

$$F = f(q_1, q_2, \dots, q_n) - \lambda \left(y - \sum_{i=1}^n p_i q_i \right)$$

and, for maximization,

$$\frac{\partial F}{\partial q_i} = f_i + \lambda p_i = 0 \quad i = 1, 2, \dots, n$$

$$\frac{\partial F}{\partial \lambda} = - \left(y - \sum_{i=1}^n p_i q_i \right) = 0$$

The demand curves for the n commodities can be obtained by solving for the q 's. The conditions

$$\frac{\partial F}{\partial q_i} = f_i + \lambda p_i = 0 \quad i = 1, 2, \dots, n$$

can be written as

$$-\frac{\partial q_i}{\partial q_j} = \frac{p_j}{p_i} \quad \text{for all } i \text{ and } j$$

That is, the rate of substitution of commodity i for commodity j must equal the price ratio p_j/p_i . The second-order conditions for maximization require that the principal minors of the bordered Hessian determinant must alternate in sign; that is,

$$\begin{bmatrix} 0 & -p_1 & -p_2 \\ -p_1 & f_{11} & f_{12} \\ -p_2 & f_{21} & f_{22} \end{bmatrix} > 0 \quad \begin{vmatrix} 0 & -p_1 & -p_2 & -p_3 \\ -p_1 & f_{11} & f_{12} & f_{13} \\ -p_2 & f_{21} & f_{22} & f_{23} \\ -p_3 & f_{31} & f_{32} & f_{33} \end{vmatrix} < 0$$

$$\dots (-1)^n \begin{vmatrix} 0 & -p_1 & -p_2 & \dots & -p_n \\ -p_1 & f_{11} & f_{12} & \dots & f_{1n} \\ -p_2 & f_{21} & f_{22} & \dots & f_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -p_n & f_{n1} & f_{n2} & \dots & f_{nn} \end{vmatrix} > 0$$

KUHN-TUCKER CONDITIONS

Consider a function of n variables $f(x_1, x_2, \dots, x_n)$ subject to the constraint $g(x_1, x_2, \dots, x_n) \leq 0$. A point $x^* = (x_1^*, x_2^*, \dots, x_n^*)$ is a local maximum of $f(x_1, x_2, \dots, x_n)$ subject to $g(x_1, x_2, \dots, x_n)$ only if there exists a nonnegative λ such that λ and $(x_1^*, x_2^*, \dots, x_n^*)$ satisfy the Kuhn-Tucker conditions:

$$h_i = \frac{\partial f}{\partial x_i} - \lambda \frac{\partial g}{\partial x_i} = 0 \quad i = 1, 2, \dots, n$$

$$\lambda g(x_1, x_2, \dots, x_n) = 0$$

$$g(x_1, x_2, \dots, x_n) \leq 0$$

These conditions are also sufficient if $f(x_1, x_2, \dots, x_n)$ is concave and the constraint is concave. Since a maximum point of $f(x_1, x_2, \dots, x_n)$ is a minimum point of $-f(x_1, x_2, \dots, x_n)$, this result is also applicable when a convex function is minimized subject to a convex constraint.

NOTE: A function $f(x_1, x_2, \dots, x_n)$ is convex in a region if for any two points $(\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n)$ and $(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$,

$$f[(1-t)\tilde{x}_1 + t\bar{x}_1, \dots, (1-t)\tilde{x}_n + t\bar{x}_n] \leq (1-t)f(\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n) + tf(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$$

The function is strictly convex if \leq can be replaced by $<$; the function is concave if \leq can be replaced by \geq and strictly concave if \leq can be replaced by $>$.

The method of Lagrange multipliers can be modified to determine the maximum or minimum of a function of n variables subject to one inequality constraint in a manner similar to the modification for a function of two variables subject to one inequality constraint. Assume that the inequality constraint holds as an equality constraint and obtain the maximum (or minimum) using the method of Lagrange multipliers. Then, if $\lambda > 0$, this maximum (or minimum) is also the maximum (or minimum) subject to the inequality constraint; if $\lambda \leq 0$, the maximum (or minimum) determined without regard to the constraint satisfies the constraint and is thus also the constrained maximum (or minimum).

EXAMPLE

Find the minimum of $f(x_1, x_2, x_3) = x_1 x_2 + x_1 x_3 + x_2 x_3$ subject to the constraint $x_1 x_2 x_3 \geq 125$.

Assuming the inequality constraint to hold as an equality constraint, $\lambda = \frac{2}{5}$ (see page 608). Since $\lambda > 0$, the minimum (5, 5, 5) assuming the equality constraint is also the minimum assuming the inequality constraint.

Alternatively, using the Kuhn-Tucker conditions,

$$\frac{\partial f}{\partial x_1} - \lambda \frac{\partial g}{\partial x_1} = x_2 + x_3 - \lambda x_2 x_3 = 0$$

$$\frac{\partial f}{\partial x_2} - \lambda \frac{\partial g}{\partial x_2} = x_1 + x_3 - \lambda x_1 x_3 = 0$$

$$\frac{\partial f}{\partial x_3} - \lambda \frac{\partial g}{\partial x_3} = x_1 + x_2 - \lambda x_1 x_2 = 0$$

$$\lambda(x_1 x_2 x_3 - 125) = 0$$

$$x_1 x_2 x_3 \geq 125$$

Either $\lambda = 0$ or $x_1 x_2 x_3 - 125 = 0$. If $\lambda = 0$,

$$x_2 + x_3 = 0$$

$$x_1 + x_3 = 0$$

$$x_1 + x_2 = 0$$

and $x_1 = x_2 = x_3 = 0$; but then $x_1 x_2 x_3 \geq 125$ is not satisfied. If $x_1 x_2 x_3 - 125 = 0$,

$$\frac{x_2 + x_3}{x_2 x_3} = \frac{x_1 + x_3}{x_1 x_3} = \frac{x_1 + x_2}{x_1 x_2}$$

and $x_1 = x_2 = x_3 = 5$; this solution satisfies the Kuhn-Tucker conditions. Thus, as above, the minimum of $f(x_1, x_2, x_3) = x_1 x_2 + x_1 x_3 + x_2 x_3$ subject to $x_1 x_2 x_3 \geq 125$ is $x_1 = x_2 = x_3 = 5$.

EXAMPLE

Find the maximum of $f(x, y, z) = -x^2 - 2y^2 - z^2 + xy + z$ subject to the constraint $x + y + z \leq 35$.

Assuming the inequality constraint to hold as an equality constraint, $\lambda = -21$ (see page 609). Since $\lambda < 0$, the unconstrained maximum is also the maximum subject to the inequality constraint.

$$\frac{\partial f}{\partial x} = -2x + y \quad \frac{\partial^2 f}{\partial x^2} = -2 \quad \frac{\partial^2 f}{\partial x \partial y} = 1$$

$$\frac{\partial f}{\partial y} = -4y + x \quad \frac{\partial^2 f}{\partial y^2} = -4 \quad \frac{\partial^2 f}{\partial x \partial z} = 0$$

$$\frac{\partial f}{\partial z} = -2z + 1 \quad \frac{\partial^2 f}{\partial z^2} = -2 \quad \frac{\partial^2 f}{\partial y \partial z} = 0$$

If $\frac{\partial f}{\partial x} = 0$, $\frac{\partial f}{\partial y} = 0$, $\frac{\partial f}{\partial z} = 0$, then

$$2x - y = 0$$

$$x - 4y = 0$$

and thus

$$x = y = 0$$

$$z = \frac{1}{2}$$

$$\Delta_1 = -2$$

$$\Delta_2 = \begin{vmatrix} -2 & 1 \\ 1 & -4 \end{vmatrix} = 8 - 1 = 7$$

$$\Delta_3 = \begin{vmatrix} -2 & 1 & 0 \\ 1 & -4 & 0 \\ 0 & 0 & -2 \end{vmatrix} = -16 - (-2) = -14$$

$\Delta_1 < 0$, $\Delta_2 > 0$, $\Delta_3 < 0$ so the point $(0, 0, \frac{1}{2})$ is a maximum.

Alternatively, using the Kuhn-Tucker conditions,

$$\frac{\partial f}{\partial x} - \lambda \frac{\partial g}{\partial x} = -2x + y - \lambda = 0$$

$$\frac{\partial f}{\partial y} - \lambda \frac{\partial g}{\partial y} = -4y + x - \lambda = 0$$

$$\frac{\partial f}{\partial z} - \lambda \frac{\partial g}{\partial z} = -2z + 1 - \lambda = 0$$

$$\lambda(x + y + z - 35) = 0 \quad x + y + z \leq 35$$

Either $\lambda = 0$ or $x + y + z - 35 = 0$. If $\lambda = 0$,

$$x = y = 0 \quad z = \frac{1}{2}$$

and $x + y + z \leq 35$ is satisfied. $f(0, 0, \frac{1}{2}) = \frac{1}{4}$. If $x + y + z = 35$, solving the first three equations,

$$-2x + y - \lambda = 0$$

$$x - 4y - \lambda = 0$$

$$2z + \lambda = 1$$

From the first two equations,

$$3x - 5y = 0$$

and from the second two equations,

$$x - 4y + 2z = 1$$

and, since $x + y + z = 35$,

$$x + 6y = 69$$

$$3x - 5y = 0$$

$$23y = 207$$

$$y = 9$$

$$x = 15$$

$$z = 11$$

The solution $x = 15$, $y = 9$, $z = 11$ satisfies $-2x + y - \lambda = 0$, $x - 4y - \lambda = 0$ and $2z + \lambda = 1$ if $\lambda = -21$. $f(15, 9, 11) = -362$, which is less than $f(0, 0, \frac{1}{2})$. Thus, as above, the maximum of $f(x, y, z) = -x^2 - 2y^2 - z^2 + xy + z$ subject to the constraint $x + y + z \leq 35$ is $x = y = 0$, $z = \frac{1}{2}$.

PROBLEMS

1. Determine the values of x_1 , x_2 , and x_3 that minimize the function

$$f(x_1, x_2, x_3) = x_1^2 + 4x_2^2 + x_3^2 - 4x_1x_2 - 6x_3$$

subject to the constraint $x_1 + x_2 + x_3 = 15$.

2. Determine the values of x_1 , x_2 , and x_3 that maximize the function

$$f(x_1, x_2, x_3) = 6x_1x_2 + 4x_2x_3 + 6x_2 - 3x_3^2 - x_2^2$$

subject to the constraint $x_1 + 2x_2 + x_3 = 75$.

3. Determine the values of x_1 , x_2 , and x_3 that minimize the function

$$f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 + x_1x_2x_3$$

subject to the constraint $x_1x_2x_3 = 1000$.

4. Determine the values of x_1 , x_2 , and x_3 that minimize the function

$$f(x_1, x_2, x_3) = 5x_1^2 + 10x_2^2 + x_3^2 - 4x_1x_2 - 2x_1x_3 - 36x_2$$

subject to the constraint $x_1 + x_2 + x_3 = 3$.

5. Determine the values of x_1 , x_2 , and x_3 that minimize the function

$$f(x_1, x_2, x_3) = x_1^2 + 4x_2^2 + x_3^2 - 4x_1x_2 - 6x_3$$

subject to the constraint $x_1 + x_2 + x_3 \geq 15$ (see Problem 1).

6. Determine the values of x_1 , x_2 , and x_3 that maximize the function

$$f(x_1, x_2, x_3) = 6x_1x_2 + 4x_2x_3 + 6x_2 - 3x_3^2 - x_2^2$$

subject to the constraint $x_1 + 2x_2 + x_3 \leq 75$ (see Problem 2).

7. Determine the values of x_1 , x_2 , and x_3 that minimize the function

$$f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 + x_1x_2x_3$$

subject to the constraint $x_1x_2x_3 \leq 1000$ (see Problem 3).

8. Determine the values of x_1 , x_2 , and x_3 that minimize the function

$$f(x_1, x_2, x_3) = 5x_1^2 + 10x_2^2 + x_3^2 - 4x_1x_2 - 2x_1x_3 - 36x_2$$

subject to the constraint $x_1 + x_2 + x_3 \geq 3$ (see Problem 4).

Answers to Odd-Numbered Problems

1. $x_1 = 8$, $x_2 = 4$, $x_3 = 3$

5. $x_1 = 8$, $x_2 = 4$, $x_3 = 3$

3. $x_1 = x_2 = x_3 = 10$

7. no finite maximum

8.3 Input-Output Analysis

Input-output analysis traces the flow of production in order to study the effects of a change in the demand for final goods on production of primary, intermediate, and finished goods. The primary purpose of input-output analysis is to calculate the output levels in various industries that would be required by particular levels of final goods demand. Input-output analysis can also be applied in studying sectors of the economy, either using a closed model where the output for a sector is equal to the sum of its inputs in the other sectors or using an open model which includes final demand in addition to demands from other sectors.

Assume that an economy is divided into n industries and that each industry produces only one type of output. Industries are usually related in the sense that each must use some of the others' products in order to operate. In addition, an economy must usually produce some finished products for final demand. Input-output analysis determines the production of each of the industries if final demand changes, assuming the structure of the economy does not change.

It is convenient to tabulate the data for input-output analysis as shown in Table 8.1, where b_{ij} is the dollar amount of the products of industry i used by industry j , h_i is

Table 8.1

Producer	User				Final demand	Total output
	1	2	...	n		
1	b_{11}	b_{12}	...	b_{1n}	h_1	x_1
2	b_{21}	b_{22}	...	b_{2n}	h_2	x_2
\vdots	\vdots	\vdots		\vdots	\vdots	\vdots
n	b_{n1}	b_{n2}	...	b_{nn}	h_n	x_n

the final demand for the products of industry i , and $x_i = \sum_{j=1}^n b_{ij} + h_i$ is the total output of industry i .

The structure of the economy can be described by the technological matrix

$$\mathbf{A} = (a_{ij})$$

where $a_{ij} = (b_{ij}/x_j)$ = dollar value of the output of industry i that industry j must purchase to produce one dollar's worth of its own products. Note that this definition of a_{ij} assumes that the purchases of intermediate product of an industry are proportional to the level of output of the industry. This assumption of constant proportionality between inputs and outputs is standard for input-output analysis.

The i th industry must produce outputs

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \quad \text{for } i = 1, 2, \dots, n$$

in order to supply the needs of all industries. The interindustry demand vector \mathbf{X} can thus be written \mathbf{AX} , where

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \quad \text{and} \quad \mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

the production of the economy must be adjusted to fulfill both interindustry needs and final demand. If the final demand vector is

$$\mathbf{H} = \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix} \quad h_i \geq 0 \quad i = 1, 2, \dots, n$$

this requirement can be written

$$\mathbf{X} = \mathbf{AX} + \mathbf{H}$$

Thus

$$[\mathbf{I} - \mathbf{A}]\mathbf{X} = \mathbf{H}$$

and

$$\mathbf{X} = [\mathbf{I} - \mathbf{A}]^{-1}\mathbf{H}$$

the matrix $\mathbf{I} - \mathbf{A}$ is known as the *Leontief matrix*.

There are a number of mathematical and practical problems in input-output analysis. For example, the technological matrix must be such that each of the x_i 's is

nonnegative; otherwise, the solution is not economically meaningful. There are also problems concerning industry classification and stability of the technological matrix.

Given a set of positive final demands, consider the problem of determining the conditions under which there is a unique set of positive production levels consistent with the set of final demands. First, since $[\mathbf{I} - \mathbf{A}]^{-1}$ must exist, $|\mathbf{I} - \mathbf{A}|$ must be nonzero. In addition, each element of $[\mathbf{I} - \mathbf{A}]^{-1}$ must be nonnegative, since otherwise an increase in final demand would result in a decrease in production at some point in the production process. It can be shown that, in order for positive gross-output levels to be associated with any given set of positive demands, the following conditions are required

$$1 > a_{ij} \geq 0 \quad \text{all } i \text{ and } j$$

$$|\mathbf{I} - \mathbf{A}| > 0$$

Since any k -industry subset of the n industries must be able to satisfy interindustry demands with some surplus to satisfy demands external to the k industries, all principal minors of the n industry Leontief determinant

$$|\mathbf{I} - \mathbf{A}| = \begin{vmatrix} 1 - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & 1 - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & 1 - a_{nn} \end{vmatrix}$$

must be positive.

EXAMPLE

Consider a very simple hypothetical economy of two industries, A and B , represented in Table 8.2, where data are in millions of dollars of products.

Determine the output vector of the economy if final demand changes to 200 for A and 100 for B .

$$\mathbf{A} = \begin{bmatrix} \frac{1}{2} & \frac{7}{16} \\ \frac{8}{25} & \frac{9}{20} \end{bmatrix}$$

$$\mathbf{I} - \mathbf{A} = \begin{bmatrix} \frac{1}{2} & -\frac{7}{16} \\ -\frac{8}{25} & \frac{11}{20} \end{bmatrix}$$

$$[\mathbf{I} - \mathbf{A}]^{-1} = \frac{200}{27} \begin{bmatrix} \frac{11}{20} & \frac{7}{16} \\ \frac{8}{25} & \frac{1}{2} \end{bmatrix}$$

$$\text{NOTE: } |\mathbf{I} - \mathbf{A}| = \frac{1}{2} \cdot \frac{11}{20} - \frac{7}{16} \cdot \frac{8}{25} = \frac{27}{200}.$$

As must be the case,

$$\mathbf{X} = [\mathbf{I} - \mathbf{A}]^{-1} \mathbf{H} = \frac{200}{27} \begin{bmatrix} \frac{11}{20} & \frac{7}{16} \\ \frac{8}{25} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 150 \\ 120 \end{bmatrix} = \begin{bmatrix} 1000 \\ 800 \end{bmatrix}$$

Table 8.2

Producer	User		Final demand	Total output
	A	B		
A	500	350	150	1000
B	320	360	120	800

And, if $\mathbf{H} = \begin{pmatrix} 200 \\ 100 \end{pmatrix}$, then

$$\mathbf{X} = [\mathbf{I} - \mathbf{A}]^{-1} \mathbf{H} = \frac{200}{27} \begin{bmatrix} \frac{11}{20} & \frac{7}{16} \\ \frac{8}{25} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 200 \\ 100 \end{bmatrix} = \begin{bmatrix} 1138\frac{8}{9} \\ 844\frac{4}{9} \end{bmatrix}$$

That is, industry A must have an output of $1138\frac{8}{9}$ and industry B must have an output of $844\frac{4}{9}$, where output is given in millions of dollars of products.

EXAMPLE

Consider a very simple hypothetical economy of three industries, A , B , and C , represented in Table 8.3, where data are in millions of dollars of products.

Determine the output vector of the economy if final demand changes to:

- (a) 50 for A , 10 for B , and 100 for C
- (b) 100 for A , 20 for B , and 60 for C
- (c) 80 for A , 100 for B , and 120 for C .

$$\mathbf{A} = \begin{bmatrix} \frac{1}{6} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

$$[\mathbf{I} - \mathbf{A}] = \begin{bmatrix} \frac{5}{6} & -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{3}{4} & -\frac{1}{3} \\ -\frac{1}{2} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

$$[\mathbf{I} - \mathbf{A}]^{-1} = \frac{864}{109} \begin{bmatrix} \frac{7}{18} & \frac{1}{4} & \frac{13}{48} \\ \frac{1}{3} & \frac{31}{72} & \frac{49}{144} \\ \frac{11}{24} & \frac{29}{72} & \frac{9}{16} \end{bmatrix}$$

$$\begin{aligned} \text{NOTE: } |\mathbf{I} - \mathbf{A}| &= \frac{5}{6} \cdot \frac{3}{4} \cdot \frac{2}{3} - \frac{1}{4} \cdot \frac{1}{3} \cdot \frac{1}{4} - \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{1}{4} \\ &\quad - \left(\frac{1}{4} \cdot \frac{3}{4} \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{5}{6} + \frac{2}{3} \cdot \frac{1}{4} \cdot \frac{1}{4} \right) \\ &= \frac{109}{864} \end{aligned}$$

$$c_{11} = (-1)^{1+1} \begin{vmatrix} \frac{3}{4} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{vmatrix} = \frac{7}{18} \quad c_{23} = (-1)^{2+3} \begin{vmatrix} \frac{5}{6} & -\frac{1}{4} \\ -\frac{1}{2} & -\frac{1}{3} \end{vmatrix} = \frac{29}{72}$$

$$c_{12} = (-1)^{1+2} \begin{vmatrix} -\frac{1}{4} & -\frac{1}{3} \\ -\frac{1}{2} & \frac{2}{3} \end{vmatrix} = \frac{1}{3} \quad c_{31} = (-1)^{3+1} \begin{vmatrix} -\frac{1}{4} & -\frac{1}{4} \\ \frac{3}{4} & -\frac{1}{3} \end{vmatrix} = \frac{13}{48}$$

$$c_{13} = (-1)^{1+3} \begin{vmatrix} -\frac{1}{4} & \frac{3}{4} \\ -\frac{1}{2} & -\frac{1}{3} \end{vmatrix} = \frac{11}{24} \quad c_{32} = (-1)^{3+2} \begin{vmatrix} \frac{5}{6} & -\frac{1}{4} \\ -\frac{1}{4} & -\frac{1}{3} \end{vmatrix} = \frac{49}{144}$$

$$c_{21} = (-1)^{2+1} \begin{vmatrix} -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{3} & \frac{2}{3} \end{vmatrix} = \frac{1}{4} \quad c_{33} = (-1)^{3+3} \begin{vmatrix} \frac{5}{6} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{3}{4} \end{vmatrix} = \frac{9}{16}$$

$$c_{22} = (-1)^{2+2} \begin{vmatrix} \frac{5}{6} & -\frac{1}{4} \\ -\frac{1}{2} & \frac{2}{3} \end{vmatrix} = \frac{31}{72}$$

Table 8.3

Producer	User			Final demand	Total output
	A	B	C		
A	90	150	225	75	540
B	135	150	300	15	600
C	270	200	300	130	900

As must be the case,

$$\mathbf{X} = [\mathbf{I} - \mathbf{A}]^{-1}\mathbf{H} = \frac{864}{109} \begin{bmatrix} \frac{7}{18} & \frac{1}{4} & \frac{13}{48} \\ \frac{1}{3} & \frac{31}{72} & \frac{49}{144} \\ \frac{11}{24} & \frac{29}{72} & \frac{9}{16} \end{bmatrix} \begin{bmatrix} 75 \\ 15 \\ 130 \end{bmatrix} = \begin{bmatrix} 540 \\ 600 \\ 900 \end{bmatrix}$$

(a) If $\mathbf{H} = \begin{bmatrix} 50 \\ 10 \\ 100 \end{bmatrix}$,

$$\mathbf{X} = [\mathbf{I} - \mathbf{A}]^{-1}\mathbf{H} = \frac{864}{109} \begin{bmatrix} \frac{7}{18} & \frac{1}{4} & \frac{13}{48} \\ \frac{1}{3} & \frac{31}{72} & \frac{49}{144} \\ \frac{11}{24} & \frac{29}{72} & \frac{9}{16} \end{bmatrix} \begin{bmatrix} 50 \\ 10 \\ 100 \end{bmatrix}$$

$$= \begin{bmatrix} 388.62 \\ 435.96 \\ 659.45 \end{bmatrix}$$

(b) If $\mathbf{H} = \begin{bmatrix} 100 \\ 20 \\ 60 \end{bmatrix}$,

$$\mathbf{X} = [\mathbf{I} - \mathbf{A}]^{-1}\mathbf{H} = \frac{864}{109} \begin{bmatrix} \frac{7}{18} & \frac{1}{4} & \frac{13}{48} \\ \frac{1}{3} & \frac{31}{72} & \frac{49}{144} \\ \frac{11}{24} & \frac{29}{72} & \frac{9}{16} \end{bmatrix} \begin{bmatrix} 100 \\ 20 \\ 60 \end{bmatrix}$$

$$= \begin{bmatrix} 476.70 \\ 494.31 \\ 694.68 \end{bmatrix}$$

(c) If $\mathbf{H} = \begin{bmatrix} 80 \\ 100 \\ 120 \end{bmatrix}$,

$$\mathbf{X} = [\mathbf{I} - \mathbf{A}]^{-1}\mathbf{H} = \frac{864}{109} \begin{bmatrix} \frac{7}{18} & \frac{1}{4} & \frac{13}{48} \\ \frac{1}{3} & \frac{31}{72} & \frac{49}{144} \\ \frac{11}{24} & \frac{29}{72} & \frac{9}{16} \end{bmatrix} \begin{bmatrix} 80 \\ 100 \\ 120 \end{bmatrix}$$

$$= \begin{bmatrix} 702.39 \\ 876.33 \\ 1144.95 \end{bmatrix}$$

PROBLEMS

1. A simple hypothetical economy of three industries, A , B , and C , is represented in the following table (data are in millions of dollars of products):

Producer	User			Final demand	Total output
	A	B	C		
A	80	100	100	40	320
B	80	200	60	60	400
C	80	100	100	20	300

- Determine the output vector for the economy if the final demand changes to a. 120 for A , 40 for B , and 10 for C ; b. 60 for A , 60 for B , and 60 for C .
2. A simple hypothetical economy of two industries, A and B , is represented in the following table (data are in tens of millions of dollars of products):

Producer	User		Final demand	Total output
	<i>A</i>	<i>B</i>		
<i>A</i>	14	6	8	28
<i>B</i>	7	18	11	36

Determine the output vector for the economy if the final demand changes to a. 16 for *A* and 3 for *B*; b. 2 for *A* and 4 for *B*.

3. A simple hypothetical economy of two industries, *A* and *B*, is represented in the following table (data are in millions of dollars of products):

Producer	User		Final demand	Total output
	<i>A</i>	<i>B</i>		
<i>A</i>	150	240	210	600
<i>B</i>	200	120	160	480

Determine the output vector for the economy if the final demand changes to a. 100 for *A* and 200 for *B*; b. 50 for *A* and 60 for *B*.

4. A simple hypothetical economy of three industries, *A*, *B*, and *C*, is represented in the following table (data are in tens of millions of dollars of products):

Producer	User			Final demand	Total output
	<i>A</i>	<i>B</i>	<i>C</i>		
<i>A</i>	5	4	3	3	15
<i>B</i>	3	10	6	1	20
<i>C</i>	3	4	4	1	12

Determine the output vector for the economy if the final demand changes to a. 1 for *A*, 2 for *B*, and 10 for *C*; b. 2 for *A*, 3 for *B*, and 2 for *C*.

Answers to Odd-Numbered Problems

1. a. 481.74 for *A*, 469.57 for *B*, and 371.74 for *C*
b. 469.56 for *A*, 542.61 for *B*, and 469.56 for *C*
3. a. 442.11 for *A* and 463.16 for *B*
b. 170.53 for *A* and 155.79 for *B*

8.4 Linear Programming

Many problems in business and economics are concerned essentially with the allocation of limited resources—money, personnel, materials, machines, space, time—in order to maximize some measure of performance or minimize some measure of cost.

The mathematical techniques for planning such allocations are referred to as *mathematical programming*; the special case in which the measure of performance or cost is a linear function and the restrictions on the availability or utilization of resources are expressible as linear equations or inequalities is referred to as *linear programming*.

More specifically, the general linear programming problem involves maximizing or minimizing a linear function of several *primary variables*, referred to as an *objective function*, subject to a set of linear equalities or inequalities referred to as *constraints*. None of the variables may be negative. (Note, however, that a negative variable may be written as the difference of two positive variables.)

Mathematically, the linear programming problem for maximization is written

$$\text{Maximize an objective function } Z = c_1 X_1 + c_2 X_2 + \cdots + c_n X_n$$

subject to the constraints

$$a_{11} X_1 + a_{12} X_2 + \cdots + a_{1n} X_n \leq b_1$$

$$a_{21} X_1 + a_{22} X_2 + \cdots + a_{2n} X_n \leq b_2$$

$$\vdots$$

$$a_{m1} X_1 + a_{m2} X_2 + \cdots + a_{mn} X_n \leq b_m$$

$$X_j \geq 0 \quad j = 1, 2, \dots, n$$

This may be written more compactly as

$$\text{Maximize } Z = \sum_{j=1}^n c_j X_j$$

$$\text{subject to } \sum_{j=1}^n a_{ij} X_j \leq b_i \quad \text{for } i = 1, 2, \dots, m$$

$$X_j \geq 0 \quad \text{for } j = 1, 2, \dots, n$$

and, in matrix notation, as

$$\text{Maximize } Z = CX$$

subject to $AX \leq B$, $X \geq O$, where

$$C = (c_1, c_2, \dots, c_n)$$

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

$$B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Similarly, a linear programming problem involving minimization may be stated in matrix notation as

$$\text{Minimize } Z = CX$$

subject to $AX \geq B$, $X \geq 0$. These formulations are referred to as the standard forms for the problems of linear maximization and minimization, respectively, because the constraint inequalities are all expressed in the form \leq for the maximization problem and in the form \geq for the minimization problem. In order to write a problem in standard form, it may be necessary to reverse the direction of an inequality by multiplying by a negative one. For example,

$$a_{11}x_1 + a_{12}x_2 \leq b$$

can be written

$$-a_{11}x_1 - a_{12}x_2 \geq -b$$

Although the standard form may be convenient, it is not necessary for the solution of a linear programming problem. The standard form is necessary for obtaining the dual, as discussed later in this section.

It can be shown that when there are m constraints in a linear programming problem, there are at most m variables or allocations in its solution. Various methods are available for determining which combination of m of the X_j 's maximizes or minimizes the objective function Z . Linear programming problems involving no more than two primary variables can be solved geometrically; although algebraic solution is usually more efficient, even for these problems, geometric solution is intuitively appealing and helps develop a basic understanding which is applicable to more complicated problems. For this reason, geometric solution is discussed below. There are several algebraic methods for solving linear programming problems; for more complicated problems, involving many variables and many constraints, the *simplex method* is the most efficient of these. The simplex method is an iterative computational routine, based on matrix algebra, which develops successively better solutions until the optimal solution is obtained. Most computer packages include a program for simplex solution of linear programming problems. Simpler solution is discussed and illustrated in this section. The interpretation of the dual of a linear programming problem is also considered.

GEOMETRIC SOLUTION

Linear programming problems involving two primary variables are solved geometrically by graphing the inequality constraints as equalities and thus determining a polygon of feasible solutions. A solution is said to be *feasible* if it satisfies all the constraints of a linear programming problem. Once the polygon of feasible solutions has been obtained, the next step is to determine which feasible solution optimizes (maximizes or minimizes) the objective function and is thus the solution of the problem. The following theorem is used to reduce the number of feasible solutions to be checked.

THEOREM: If there is a unique solution which maximizes or minimizes a linear objective function, then that solution must correspond to a vertex (or corner) of the polygon of feasible solutions; if there is more than one solution, at least two of the solutions must correspond to adjacent vertices of the polygon of feasible solutions.

Thus the value of the objective function need be computed only for solutions corresponding to vertices of the polygon of feasible solutions in order to determine the optimal solution. (In some cases certain of these solutions are clearly inferior to others and the corresponding values of the objective function need not be computed.)

EXAMPLE

A manufacturer produces bicycles and motor scooters, each of which must be processed through two machine centers. Machine center 1 has a maximum of 120 hours available and machine center 2 has a maximum of 180 hours available. Manufacturing a bicycle requires 6 hours in machine center 1 and 3 hours in machine center 2; manufacturing a motor scooter requires 4 hours in machine center 1 and 10 hours in machine center 2. If profit is \$45 for a bicycle and \$55 for a motor scooter, determine the number of bicycles and the number of motor scooters that should be manufactured in order to maximize profit.

If X_1 = number of bicycles and X_2 = number of motor scooters, the problem can be stated as follows:

$$\text{Maximize } Z = 45X_1 + 55X_2$$

subject to $6X_1 + 4X_2 \leq 120$, $3X_1 + 10X_2 \leq 180$, and $X_1, X_2 \geq 0$. The polygon of feasible solutions is obtained by graphing the equations

$$6X_1 + 4X_2 = 120$$

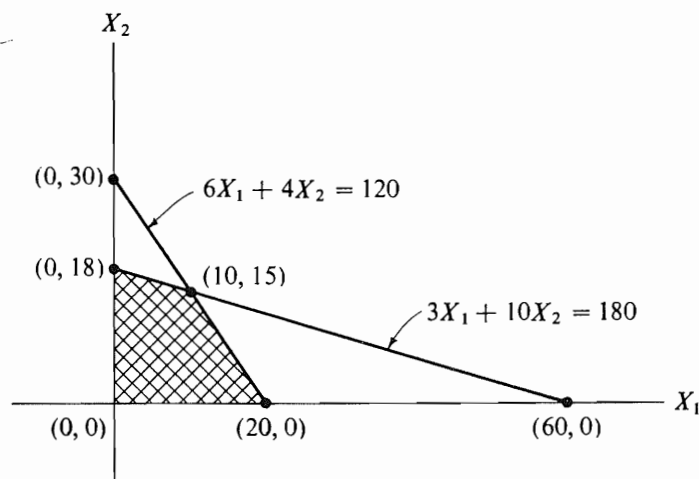
$$3X_1 + 10X_2 = 180$$

$$X_1 = 0$$

$$X_2 = 0$$

Any point lying within or on the edge of the shaded polygon in Figure 8.1 corresponds to a pair of values X_1 and X_2 which satisfy the constraints of the problem and thus are a feasible

Figure 8.1



solution. The optimal solution corresponds to one of the vertices of the polygon; the objective function thus must be evaluated for the solutions $(0, 0)$, $(0, 18)$, $(10, 15)$, and $(20, 0)$.

Solution	$Z = 45X_1 + 55X_2$
$(0, 18)$	\$ 990
$(10, 15)$	\$1275
$(20, 0)$	\$ 900

Since $(0, 0)$ is clearly inferior to $(0, 18)$, it need not be considered. Thus the optimal solution is $(10, 15)$; that is, manufacture 10 bicycles and 15 motor scooters.

EXAMPLE

A company produces two commodities in quantities X_1 and X_2 , respectively, and wishes to minimize $\text{Cost} = 2X_1 + 10X_2$ subject to the constraints

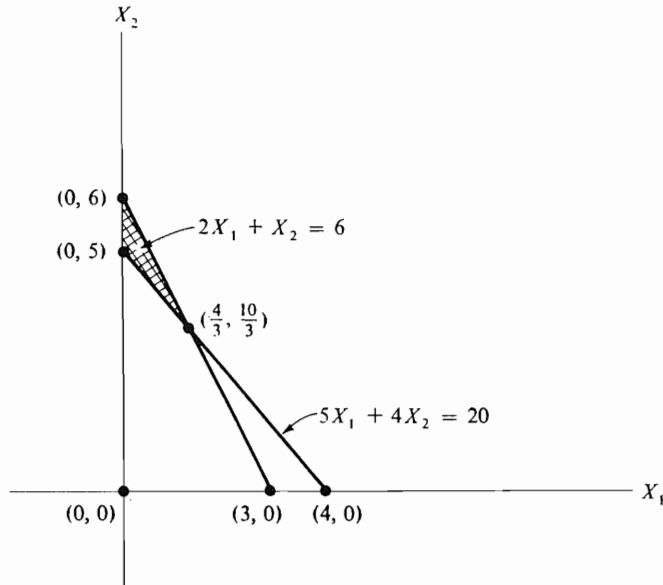
$$2X_1 + X_2 \leq 6$$

$$5X_1 + 4X_2 \geq 20$$

$$X_1, X_2 \geq 0$$

Determine the optimal quantities of each commodity to be produced and the associated cost.

Figure 8.2



The polygon of feasible solutions is obtained by graphing the equations

$$2X_1 + X_2 = 6$$

$$5X_1 + 4X_2 = 20$$

$$X_1 = 0$$

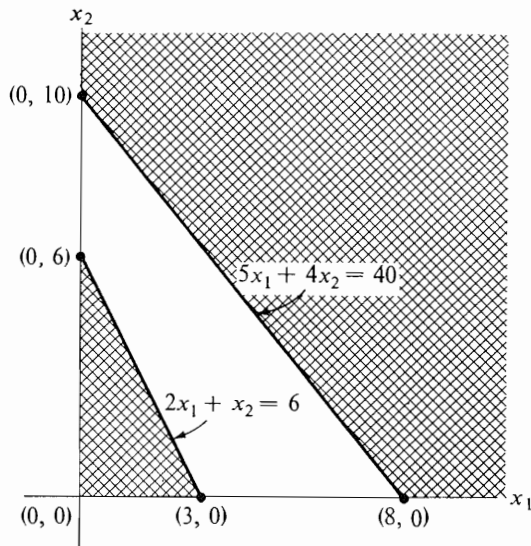
$$X_2 = 0$$

Any point lying in or on the edge of the shaded polygon in Figure 8.2 corresponds to a feasible solution. The optimal solution is one of the vertices $(0, 5)$, $(0, 6)$, and $(\frac{4}{3}, \frac{10}{3})$.

Solution	$Z = 2X_1 + 10X_2$
$(0, 5)$	\$50
$(\frac{4}{3}, \frac{10}{3})$	\$36

Note that $(0, 6)$ need not be evaluated, since clearly it is associated with a higher cost than $(0, 5)$. The optimal solution is $(\frac{4}{3}, \frac{10}{3})$; that is, produce $\frac{4}{3}$ units of the first commodity and $\frac{10}{3}$ units of the second commodity. The associated cost is \$36.

Figure 8.3



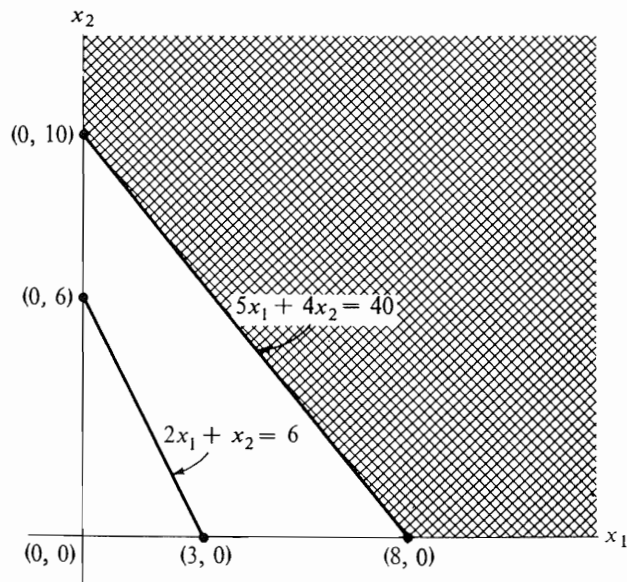
Note that the constraints determine the polygon of feasible solutions. The objective function is used in determining which of these feasible solutions is optimal. It is possible to write sets of constraints for which there is no feasible solution. For example, there is no set of values for X_1 and X_2 which satisfies the constraints

$$2X_1 + X_2 \leq 6$$

$$5X_1 + 4X_2 \geq 40$$

(See Figure 8.3.)

Figure 8.4



There are also some sets of constraints for which the polygon of feasible solutions is not bounded. For example, if the inequality sign of the first constraint above is reversed and the second constraint is unchanged,

$$2X_1 + X_2 \geq 6$$

$$5X_1 + 4X_2 \geq 40$$

the polygon of feasible solutions is unbounded. (See Figure 8.4.) In this case, the linear programming problem has a solution if it involves minimization, but not if it involves maximization.

THE SIMPLEX METHOD

The simplex method solves linear programming problems by obtaining a feasible solution and, by an iterative procedure, improving this solution until the optimal solution is obtained. The computational routine of the simplex method is based on matrix algebra and consists essentially of obtaining an inverse matrix in order to solve a set of simultaneous linear equations.

For simplex solution it is necessary to write the constraint inequalities as equalities (that is, equations) by adding positive variables, referred to as *slack variables*, to the left-hand side of the inequalities. For a maximization problem written in standard form, one slack variable is added to the left-hand side of each inequality. If the slack variable added to the i th equation is denoted by X_{n+i} , the problem can be rewritten as follows:

$$\text{Maximize } Z = c_1 X_1 + c_2 X_2 + \cdots + c_n X_n$$

subject to

$$a_{11} X_1 + a_{12} X_2 + \cdots + a_{1n} X_n + X_{n+1} = b_1$$

$$a_{21} X_1 + a_{22} X_2 + \cdots + a_{2n} X_n + X_{n+2} = b_2$$

$$\vdots$$

$$a_{m1} X_1 + a_{m2} X_2 + \cdots + a_{mn} X_n + X_{n+m} = b_m$$

$$X_j \geq 0 \quad j = 1, 2, \dots, n + m$$

In this problem there are n primary variables and m constraints. The addition of slack variables results in m equations involving $m + n$ unknowns. Note that the slack variables are omitted from the objective function or, equivalently, are given a value of zero in the objective function.

If an inequality is of the form \geq , a positive slack variable is subtracted from the left-hand side of the inequality; this slack variable is given a value of zero in the objective function. In order to obtain a feasible solution, an artificial variable is added to the left-hand side of the inequality; this artificial variable is given a value of $-M$ in the objective function for a maximization problem and a value of $+M$ in the objective function for a minimization problem, where M is very large in absolute value. The artificial variable is thus assigned a value in the objective function that precludes its appearance in the final solution. For example, if an inequality is of the form

$$a_{11} X_1 + a_{12} X_2 + \cdots + a_{1n} X_n \geq b_1$$

Table 8.4

Profits per unit			Variables in the solution			Quantities of the variables ↓ in the solution		

Table 8.5

C_j	Variables allocated	Quantity	\$45	\$55	\$0	\$0
			X_1	X_2	X_3	X_4
\$0	X_3	120	6	4	1	0
\$0	X_4	180	3	10	0	1
	Z_j	\$0	\$0	\$0	\$0	\$0
	$C_j - Z_j$		\$45	\$55	\$0	\$0

↑
Optimal column

With the necessary slack variables included, the data are arranged in a simplex tableau, and an initial feasible solution is indicated in the tableau (steps 1 through 3). The initial simplex tableau is shown in Table 8.5.

Note the following characteristics of the initial simplex tableau:

- a. Any unknown that occurs in one equation occurs (implicitly at least) in all equations; the unknowns that do not appear explicitly in an equation are considered to have zero coefficients in that equation. Thus the zeros in the X_4 column of the first row and the X_3 column of the second row correspond to those in the equations

$$6X_1 + 4X_2 + X_3 + 0X_4 = 120$$

$$3X_1 + 10X_2 + 0X_3 + X_4 = 180$$

- b. The initial tableau is constructed for the solution having zero profit. The objective function can be written

$$Z = 45X_1 + 55X_2 + 0X_3 + 0X_4$$

If $X_3 = 120$ and $X_4 = 180$, then $X_1 = X_2 = 0$ and Z is also zero. This solution is obviously not optimal, but provides a convenient initial feasible solution. The solution is then checked for optimality.

Step 4. Check for optimality: The $C_j - Z_j$ row gives the net profit from adding one unit of a variable. For a maximization problem, the presence of at least one positive number in the $C_j - Z_j$ row indicates that profit can be improved; absence of positive numbers in the $C_j - Z_j$ row indicates that profit cannot be improved, that is, the optimal solution has been obtained. For a minimization problem, the presence of at least one negative number in the $C_j - Z_j$ row indicates that cost can be decreased; absence of negative numbers in the $C_j - Z_j$ row indicates that cost cannot be decreased, that is, the optimal solution has been obtained. In this example, there are positive numbers in the $C_j - Z_j$ row of the initial tableau, so the corresponding solution is not optimal.

Step 5. Entering and departing variables: The largest contribution to profit per unit (\$55) is made by variable X_2 , so it is the entering or replacing variable. To determine the departing variable, the quantities of X_3 and X_4 in the initial tableau are divided respectively by their corresponding entries in the optimal column:

$$X_3 \text{ row: } \frac{120 \text{ hours available}}{4 \text{ hours per unit}} = 30 \text{ units of } X_2$$

$$X_4 \text{ row: } \frac{180 \text{ hours available}}{10 \text{ hours per unit}} = 18 \text{ units of } X_2$$

Table 8.6

C_j	Variables allocated	Quantity	\$45	\$55	\$0	\$0
			X_1	X_2	X_3	X_4
\$0	X_3	48	$\left(\frac{24}{5}\right)$	0	1	$-\frac{2}{5}$
\$55	X_2	18	$\left(\frac{3}{10}\right)$	1	0	$\frac{1}{10}$
	Z_j	\$990	$\$ \frac{33}{2}$	\$55	\$0	$\$ \frac{11}{2}$
	$C_j - Z_j$		$\$ \frac{57}{2}$	\$0	\$0	$\$ -\frac{11}{2}$

↑

The variable corresponding to the smaller positive ratio is the departing or replaced variable; thus X_4 is the replaced variable in the example. (Note that 18 is the largest number of units of X_2 than can be added without violating the constraints.)

Step 6. Recompute the simplex tableau: The X_2 row of the revised tableau is computed by dividing each number in the replaced row X_4 by the element in the optimal column of the replaced row, which is 10 in this example (Table 8.6).

The entries in the X_2 row are

$$18 \quad \frac{3}{10} \quad 1 \quad 0 \quad \frac{1}{10}$$

All the remaining rows corresponding to variables in the tableau are computed using the formula

$$\left(\begin{array}{c} \text{element in} \\ \text{new row} \end{array} \right) = \left(\begin{array}{c} \text{element in} \\ \text{old row} \end{array} \right) - \left(\begin{array}{c} \text{element of old row} \\ \text{in the optimal} \\ \text{column} \end{array} \right) \times \left(\begin{array}{c} \text{corresponding} \\ \text{element in} \\ \text{replacing row} \end{array} \right)$$

In this example, the only remaining variable row is X_3 ; its entries in the revised tableau are

$$120 - (4 \times 18) = 48$$

$$6 - (4 \times \frac{3}{10}) = \frac{24}{5}$$

$$4 - (4 \times 1) = 0$$

$$1 - (4 \times 0) = 1$$

$$0 - (4 \times \frac{1}{10}) = -\frac{2}{5}$$

$$Z_j \text{ is the total profit for the solution} = 0 \cdot 48 + 55 \cdot 18 = 990$$

$$\left. \begin{array}{l} Z_j \text{ for } X_1 = (0)(\frac{24}{5}) + (55)(\frac{3}{10}) = \frac{33}{2} \\ Z_j \text{ for } X_2 = (0)(0) + (55)(1) = 55 \\ Z_j \text{ for } X_3 = (0)(1) + (55)(0) = 0 \\ Z_j \text{ for } X_4 = (0)(-\frac{2}{5}) + (55)(\frac{1}{10}) = \frac{11}{2} \end{array} \right\} \begin{array}{l} \text{profit given up by adding} \\ \text{one unit of the variable} \end{array}$$

The $C_j - Z_j$ row gives the net profit per unit (profit per unit minus profit given up by adding one unit). The positive entry $\$ \frac{57}{2}$ in the revised tableau indicates that the solution can be improved by the inclusion of variable X_1 . Checking to see which variable is to be omitted (although it is obviously X_3 in this example),

$$X_3: \frac{48}{\frac{24}{5}} = 10$$

$$X_2: \frac{18}{\frac{3}{10}} = 60$$

Table 8.7

C_j	Variables allocated	Quantity	\$45	\$55	\$0	\$0
			X_1	X_2	X_3	X_4
\$45	X_1	10	1	0	$\frac{5}{24}$	$-\frac{1}{12}$
\$55	X_2	15	0	1	$-\frac{1}{16}$	$\frac{1}{8}$
	Z_j	\$1275	\$45	\$55	$\$ \frac{95}{16}$	$\$ \frac{25}{8}$
	$C_j - Z_j$		\$0	\$0	$\$ -\frac{95}{16}$	$\$ -\frac{25}{8}$

The smallest positive ratio corresponds to X_3 , and X_3 is thus the departing variable. The simplex tableau is recomputed the same way as in the preceding stage (see Table 8.7). The absence of positive entries in the $C_j - Z_j$ row indicates that the optimal solution has been obtained: Manufacture 10 bicycles and 15 motor scooters, as was also determined above using the geometric method.

Note the following characteristics of the simplex tableau:

- Slack and artificial variables are included in the constraint equations as follows: If an inequality constraint is of the form \leq , a slack variable is added and is given a value of zero in the objective function for both maximization and minimization problems. If an inequality constraint is of the form \geq , a slack variable is subtracted and is given a value of zero in the objective function for both maximization and minimization problems; in addition, an artificial variable is added and is given a value in the objective function of $-M$ for maximization problems and $+M$ for minimization problems.
- The elements of the body and identity matrices of the simplex tableau represent marginal rates of substitution between the variables in the solution and the variables heading the columns. For example, in Table 8.6, X_3 must be decreased by $\frac{24}{5}$ units if 1 unit of X_1 is added; X_3 must be increased by $\frac{2}{3}$ units if 1 unit of X_4 is added, and so forth. Note that a variable always has a marginal rate of substitution of 1 with itself and a marginal rate of substitution of 0 with a variable with which it is not substitutable.
- The body matrix of the initial tableau has become an identity matrix in the final tableau; the identity matrix of the initial tableau has become the inverse of the initial body matrix in the final tableau.
- In the final tableau, the $C_j - Z_j$ row consists of zero or negative numbers for a maximization problem and of zero or positive numbers for a minimization problem.

EXAMPLE

Consider the problem of cost minimization discussed above (see page 623). Adding the slack variables X_3 and X_4 and the artificial variable X_5 , the linear programming problem can be stated as follows.

$$\text{Minimize } Z = 2X_1 + 10X_2 + MX_5$$

subject to

$$2X_1 + X_2 + X_3 = 6$$

$$5X_1 + 4X_2 - X_4 + X_5 = 20$$

$$X_1, X_2, X_3, X_4, X_5 \geq 0$$

The initial simplex tableau is shown in Table 8.8.

Table 8.8

C_j	Variables allocated	Quantity	\$2	\$10	\$0	\$0	\$M
			X_1	X_2	X_3	X_4	X_5
\$0	X_3	6	2	1	1	0	0
\$M	X_5	20	5	4	0	-1	1
	Z_j	$20M$	$5M$	$4M$	0	$-M$	M
	$C_j - Z_j$		$$(2 - 5M)$	$$(10 - 4M)$	0	$$M$	0

↑
Optimal column

The variable to enter at the next step is X_1 and the variable to be replaced is X_3 since

$$X_3: \frac{6}{2} = 3$$

$$X_5: \frac{20}{5} = 4$$

The revised simplex tableau is shown in Table 8.9.

Table 8.9

C_j	Variables allocated	Quantity	\$2	\$10	\$0	\$0	\$M
			X_1	X_2	X_3	X_4	X_5
\$2	X_1	3	1	$\frac{1}{2}$	$\frac{1}{2}$	0	0
\$M	X_5	5	0	$\frac{3}{2}$	$-\frac{5}{2}$	-1	1
	Z_j	$$(6 + 5M)$	2	$$(1 + 3M/2)$	$$(1 - 5M/2)$	$-M$	M
	$C_j - Z_j$		0	$$(9 - 3M/2)$	$$(-1 + 5M/2)$	$$M$	0

↑
Optimal column

The variable to enter at the next step is X_2 and the variable to be replaced is X_5 since

$$X_1: \frac{3}{\frac{1}{2}} = 6$$

$$X_5: \frac{5}{\frac{3}{2}} = \frac{10}{3}$$

The revised simplex tableau is shown in Table 8.10. The absence of negative entries in the $C_j - Z_j$ row indicates that the optimal solution has been obtained: manufacture $\frac{4}{3}$ units of the first commodity and $\frac{10}{3}$ units of the second commodity at a total cost of \$36, as was also determined above by the graphical method.

Table 8.10

C_j	Variables allocated	Quantity	\$2	\$10	\$0	\$0	\$M
			X_1	X_2	X_3	X_4	X_5
\$2	X_1	$\frac{4}{3}$	1	0	$\frac{4}{3}$	$\frac{1}{3}$	$-\frac{1}{3}$
\$10	X_2	$\frac{10}{3}$	0	1	$-\frac{5}{3}$	$-\frac{2}{3}$	$\frac{2}{3}$
	Z_j	36	2	10	-14	-6	6
	$C_j - Z_j$		0	0	14	6	$$(M - 6)$

There are several variations of the standard simplex method, although none differs substantially from the one discussed above. In addition, there are some linear programming problems in which the coefficients in the constraints have special forms and special cases of the simplex method, which require less computation, can be used. The simplex method must be modified if the solution variables are required to be integer-valued.

THE DUAL OF A LINEAR PROGRAMMING PROBLEM

Corresponding to every linear programming problem is a second linear programming problem referred to as the *dual*. When the initial or *primal* problem involves maximization (minimization) of an objective function, the dual problem involves minimization (maximization) of an objective function. The number of variables in the dual problem is equal to the number of constraints in the primal problem and the number of constraints in the dual problem is equal to the number of variables in the primal problem.

For the linear programming problem stated above (page 620), the dual can be stated as follows:

$$\text{Minimize } Z_{\text{dual}} = b_1 Y_1 + b_2 Y_2 + \cdots + b_m Y_m$$

subject to

$$a_{11} Y_1 + a_{21} Y_2 + \cdots + a_{m1} Y_m \geq c_1$$

$$a_{12} Y_1 + a_{22} Y_2 + \cdots + a_{m2} Y_m \geq c_2$$

$$\vdots$$

$$a_{1n} Y_1 + a_{2n} Y_2 + \cdots + a_{mn} Y_m \geq c_m$$

$$Y_j \geq 0 \quad j = 1, 2, \dots, m$$

Thus if a linear programming problem, written in standard form, is

$$\text{Maximize } CX$$

subject to

$$AX \leq B$$

$$X \geq 0$$

then its dual, written in standard form, is

$$\text{Minimize } BY$$

subject to

$$A'Y \geq C$$

$$Y \geq 0$$

and, similarly, if a linear programming problem, written in standard form, is

$$\text{Minimize } CX$$

subject to

$$AX \geq B$$

$$X \geq 0$$

then its dual, written in standard form, is

Maximize \mathbf{BY}

subject to

$$\mathbf{A}'\mathbf{Y} \leq \mathbf{C}$$

$$\mathbf{Y} \geq \mathbf{0}$$

The correspondence between the primal and dual problems can be summarized as follows:

1. The constants in the constraints of the primal are the coefficients in the objective function of the dual.
2. The coefficients in the objective function of the primal are the constants in the constraints of the dual.
3. The inequality signs in the constraints of the primal and dual are reversed.
4. The matrices of coefficients in the constraints of the primal and dual are transposes of each other; for example, the i th row of coefficients in the constraints of the primal is the i th column of coefficients in the constraints of the dual.

EXAMPLE

Write the dual of each of the following linear programming problems.

(a) Min $5X_1 + 9X_2$

subject to

$$3X_1 + 2X_2 \leq 6$$

$$5X_1 + X_2 \geq 10$$

$$X_1 + 10X_2 \geq 9$$

$$X_1 \geq 0 \quad X_2 \geq 0$$

Writing the problem in standard form,

$$\text{Min } 5X_1 + 9X_2$$

subject to

$$-3X_1 - 2X_2 \geq -6$$

$$5X_1 + X_2 \geq 10$$

$$X_1 + 10X_2 \geq 9$$

$$X_1 \geq 0 \quad X_2 \geq 0$$

and its dual is

$$\text{Max } -6Y_1 + 10Y_2 + 9Y_3$$

subject to

$$-3Y_1 + 5Y_2 + Y_3 \leq 5$$

$$-2Y_1 + Y_2 + 10Y_3 \leq 9$$

$$Y_1 \geq 0 \quad Y_2 \geq 0 \quad Y_3 \geq 0$$

(b) Max $X_1 + 9X_2 + 15X_3$

subject to

$$3X_1 + 2X_2 \geq 11$$

$$X_1 + X_2 + X_3 \geq 15$$

$$8X_2 + 7X_3 \leq 25$$

$$X_1 \geq 0 \quad X_2 \geq 0 \quad X_3 \geq 0$$

Writing the problem in standard form,

$$\text{Max } X_1 + 9X_2 + 15X_3$$

subject to

$$-3X_1 - 2X_2 \leq -11$$

$$-X_1 - X_2 - X_3 \leq -15$$

$$8X_2 + 7X_3 \leq 25$$

$$X_1 \geq 0 \quad X_2 \geq 0 \quad X_3 \geq 0$$

and its dual is

$$\text{Min } -11Y_1 - 15Y_2 + 25Y_3$$

subject to

$$-3Y_1 - Y_2 \geq 1$$

$$-2Y_1 - Y_2 + 8Y_3 \geq 9$$

$$-Y_2 + 7Y_3 \geq 15$$

$$Y_1 \geq 0 \quad Y_2 \geq 0 \quad Y_3 \geq 0$$

The solution of the dual problem also provides the solution of the primal problem. The number of iterations required for simplex solution depends in part on the number of rows of variables in the simplex tableau; thus, if $m < n$, solution of the dual problem usually requires less computation and is therefore preferable.

The correspondence between the primal and dual solutions can be summarized as follows:

1. The value of the objective function is the same for the primal and dual solutions.
2. The criteria for the primary variables of the primal are the solutions for the slack variables of the dual.
3. The criteria for the slack variables of the primal are the solutions for the primary variables of the dual.
4. The solutions for the primary variables of the primal are the negative values of the criteria for the slack variables of the dual.
5. The solutions for the slack variables of the primal are the negative values of the criteria for the primary variables of the dual.

EXAMPLE

For the maximization problem on page 622,

$$\text{Max } Z = 45X_1 + 55X_2$$

Table 8.11

			\$120	\$180	\$0	\$M	\$0	\$M
C_j	Variables allocated	Quantity	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6
\$120	Y_1	$\frac{95}{16}$	1	0	$-\frac{5}{24}$	$\frac{5}{24}$	$\frac{1}{16}$	$-\frac{1}{16}$
\$180	Y_2	$\frac{25}{8}$	0	1	$\frac{1}{12}$	$-\frac{1}{12}$	$-\frac{1}{8}$	$\frac{1}{8}$
	Z_j	\$1275	\$120	\$180	\$ - 10	\$10	\$ - 15	\$15
	$C_j - Z_j$		\$0	\$0	\$10	$\$(M - 10)$	\$15	$\$(M - 15)$

subject to

$$6X_1 + 4X_2 \leq 120$$

$$3X_1 + 10X_2 \leq 180$$

$$X_1, X_2 \geq 0$$

the dual is given by

$$\text{Min } 120Y_1 + 180Y_2$$

subject to

$$6Y_1 + 3Y_2 \geq 45$$

$$4Y_1 + 10Y_2 \geq 55$$

$$Y_1, Y_2 \geq 0$$

With slack and artificial variables, the dual problem is

$$\text{Min } 120Y_1 + 180Y_2 + MY_4 + MY_6$$

subject to

$$6Y_1 + 3Y_2 - Y_3 + Y_4 = 45$$

$$4Y_1 + 10Y_2 - Y_5 + Y_6 = 55$$

$$Y_1, Y_2, Y_3, Y_4, Y_5, Y_6 \geq 0$$

The final tableau of the simplex solution of the dual is given in Table 8.11. Note the correspondence between Tables 8.7 and 8.11 with respect to the properties listed above.

EXAMPLE

For the minimization problem on page 623,

$$\text{Min } 2X_1 + 10X_2$$

subject to

$$2X_1 + X_2 \leq 6$$

$$5X_1 + 4X_2 \geq 20$$

$$X_1, X_2 \geq 0$$

the dual is given by

$$\text{Max } -6Y_1 + 20Y_2$$

Table 8.12

C_j	Variables allocated	Quantity	\$ - 6	\$20	\$0	\$0
			Y_1	Y_2	Y_3	Y_4
\$20	Y_2	6	0	1	$-\frac{1}{3}$	$\frac{2}{3}$
\$ - 6	Y_1	14	1	0	$-\frac{4}{3}$	$\frac{5}{3}$
	Z_j	\$36	\$ - 6	\$20	\$ $\frac{4}{3}$	\$ $\frac{10}{3}$
	$C_j - Z_j$		\$0	\$0	\$ $-\frac{4}{3}$	\$ $-\frac{10}{3}$

subject to

$$-2Y_1 + 5Y_2 \leq 2$$

$$-Y_1 + 4Y_2 \leq 10$$

$$Y_1, Y_2 \geq 0$$

With slack and artificial variables, the dual problem is

$$\text{Max } -6Y_1 + 20Y_2$$

subject to

$$-2Y_1 + 5Y_2 + Y_3 = 2$$

$$-Y_1 + 4Y_2 + Y_4 = 10$$

$$Y_1, Y_2, Y_3, Y_4 \geq 0$$

The final tableau of the simplex solution of the dual is given in Table 8.12. Note the correspondence between Tables 8.10 and 8.12 with respect to the properties listed above.

Thus the dual problem is completely symmetric to the primal problem and the solution of one problem provides the solution of the other problem. Since the number of iterations required for simplex solution depends in part on the number of rows in the simplex tableau, the dual can provide a simpler computational problem than the primal when the simplex tableau of the dual has fewer rows than the simplex tableau of the primal.

The solution of the dual provides *implicit*, *imputed*, or *shadow prices* that can be used as the basis for efficient allocation of resources. These shadow prices are also used in sensitivity analysis, which determines the effect on the optimal solution if constraints or objective function coefficients change. These analyses are discussed in economics and are only mentioned here to indicate some of the applications of the dual.

PROBLEMS

1. A company manufactures two types of novelty souvenirs made of plywood. Souvenirs of type A require 5 minutes each for cutting and 10 minutes each for assembling; souvenirs of type B require 8 minutes each for cutting and 8 minutes each for assembling. There are 3 hours and 20 minutes available for cutting and 4 hours available for assembling. The profit is 50¢ for each type A souvenir and 60¢ for each type B souvenir. How many souvenirs of each type should the company manufacture in order to maximize profit?

2. Determine the values of x and y that maximize the function

$$f(x, y) = 0.4x + 0.9y$$

subject to the constraints

$$5x + 3y \leq 30 \quad 7x + 2y \leq 28 \quad x \geq 0 \quad y \geq 0$$

3. Determine the values of x_1 and x_2 that maximize the function

$$f(x, x_2) = 3x_1 + 5x_2$$

subject to the constraints

$$x_1 + 2x_2 \leq 10 \quad 3x_1 + x_2 \leq 10 \quad x_1 \geq 0 \quad x_2 \geq 0$$

4. Determine the values of x_1 and x_2 that maximize the function

$$f(x_1, x_2) = 3x_1 + 6x_2$$

subject to the constraints

$$2x_1 + 5x_2 \leq 20 \quad 2x_1 + 2x_2 \leq 10 \quad x_1 \geq 0 \quad x_2 \geq 0$$

5. Determine the values of x and y that maximize the function

$$f(x, y) = 3x + 2y$$

subject to the constraints

$$x + y \leq 7 \quad 2x + 3y \leq 16 \quad x \geq 0 \quad y \geq 0$$

6. A manufacturer of a line of hair shampoos is planning a production run for products A (for dry hair) and B (for oily hair). There are sufficient ingredients on hand for 60,000 bottles of each shampoo but there are only 60,000 bottles into which either shampoo can be put. It takes 4 hours to prepare enough shampoo to fill 1000 bottles of shampoo A and 3 hours to prepare enough shampoo to fill 1000 bottles of shampoo B ; there are 200 hours available for preparation. Profit is 9 cents a bottle for shampoo A and 7 cents a bottle for shampoo B . How should production be scheduled to maximize profit?
7. The Tr-Ply Company manufactures two types of heavy-duty wood crates. The profit on each crate of type 1 is \$9 and the profit on each crate of type 2 is \$12. Each crate must go through two production lines; a total of 10 hours are available on production line A and a total of 12 hours are available on production line B . Each crate of type 1 requires 2 hours on line A and 4 hours on line B . Each crate of type 2 requires 5 hours on line A and 3 hours on line B . Determine the number of crates of each type that should be produced in order to maximize the associated profit.
8. The UP-DOWN Ladder Company can manufacture three types of ladders; profit is \$5 per ladder of type 1, \$7 per ladder of type 2, and \$8 per ladder of type 3. Each ladder must be processed through three centers according to the following requirements:

	Center 1 (min)	Center 2 (min)	Center 3 (min)
Type 1	4	5	6
Type 2	5	7	9
Type 3	6	7	7
Total available	80	100	120

In order to maximize profit, determine the number of ladders of each type that should be manufactured.

9. The Capsize Yachting Company produces two models of racing boats. Their profit is \$520 for Model 1 and \$450 for Model 2. Model 1 requires 40 hours for cutting and assembling

and 24 hours for finishing. Model 2 requires 25 hours for cutting and assembling and 30 hours for finishing. There are 400 hours available for cutting and assembling and 360 hours available for finishing. Determine the optimal number of each model for the company to produce and the resulting profit.

10. The Duoply Company can make three products; their profit is \$10 per unit for Product 1, \$14 per unit for Product 2, and \$15 per unit for Product 3. Each of these products is manufactured using three raw materials according to the following requirements:

	Material 1	Material 2	Material 3
Product 1	3 lb/unit	4 lb/unit	5 lb/unit
Product 2	5 lb/unit	7 lb/unit	7 lb/unit
Product 3	4 lb/unit	8 lb/unit	6 lb/unit
Total lb available	220	280	320

Determine the optimal product mix.

11. College Publishers, Inc., plans to use one section of its plant to produce two textbooks. The profit is \$2 for Textbook 1 and \$3 for Textbook 2. Textbook 1 requires 4 hours for printing and 6 hours for binding. Textbook 2 requires 5 hours for printing and 3 hours for binding. There are 200 hours available for printing and 210 hours available for binding. Determine the optimal number of each textbook for the publishers to produce and their resulting profit.
12. Obtain the solution for the following linear programming problem.

$$\text{Maximize } 6X_1 + 3X_2$$

subject to

$$4X_1 + X_2 \leq 12$$

$$2X_1 + 2X_2 \leq 10$$

$$2X_1 + 4X_2 \geq 8$$

$$X_1, X_2 \geq 0$$

13. Obtain the solution for the following linear programming problem.

$$\text{Minimize } 12X_1 - 5X_2$$

subject to

$$X_1 - 2X_2 \geq 3$$

$$X_1 - X_2 \geq 4$$

$$X_1 \geq 1$$

$$X_1, X_2 \geq 0$$

Write the duals of the following linear programming problems.

14. Max $X_1 + X_2 + 5X_3$

$$\text{subject to } 4X_1 + 3X_2 + X_3 \leq 10$$

$$2X_1 + 10X_2 + 3X_3 \geq 15$$

$$X_1, X_2, X_3 \geq 0$$

15. Min $3X_1 + 5X_2 + X_3$

subject to $X_1 + X_2 + X_3 \geq 6$

$$3X_1 + 8X_2 + 9X_3 \leq 50$$

$$6X_1 + 7X_3 \geq 12$$

$$12X_2 + 4X_3 \geq 15$$

$$X_1, X_2, X_3 \geq 0$$

16. Max $5X_1 - 6X_2 + 10X_3$

subject to $X_1 + X_2 + X_3 \geq 15$

$$2X_1 + 3X_2 + 4X_3 \leq 35$$

$$3X_1 - 4X_2 + 6X_3 \leq 30$$

$$X_1 - X_2 \geq 0$$

$$X_1, X_2, X_3 \geq 0$$

17. Min $2X_1 + 10X_2$

subject to $X_1 + X_2 + X_3 \geq 5$

$$3X_1 - 2X_2 + 6X_3 \leq 20$$

$$X_2 + 3X_3 \geq 10$$

$$X_1, X_2, X_3 \geq 0$$

18. Max $8X_1 - 3X_2$

subject to $9X_1 + 2X_2 \leq 20$

$$10X_1 + 3X_2 \leq 32$$

$$X_1, X_2 \geq 0$$

19. Max $5X_1 + 3X_2 + 14X_3$

subject to $2X_1 + X_2 + 3X_3 \leq 14$

$$X_1 + 3X_2 + 2X_3 \leq 15$$

$$X_1 + X_2 + X_3 \geq 8$$

$$X_1, X_2, X_3 \geq 0$$

20. Min $6X_1 - 3X_2 + 4X_3$

subject to $3X_1 + 6X_2 + 2X_3 \leq 30$

$$5X_1 + X_2 + 6X_3 \geq 25$$

$$X_1, X_2, X_3 \geq 0$$

Answers to Odd-Numbered Problems

1. 8 type A and 20 type B

3. $x_1 = 2$, $x_2 = 4$

5. $x = 7$, $y = 0$

7. $\frac{15}{7}$ type 1, $\frac{8}{7}$ type 2, profit \$33

9. 5 Model 1, 8 Model 2, profit \$6200

11. 0 Textbook 1, 40 Textbook 2, profit \$120

13. $X_1 = 4$, $X_2 = 0$, min = 48

15. Max $6Y_1 - 50Y_2 + 12Y_3 + 15Y_4$

subject to $Y_1 - 3Y_2 + 6Y_3 \leq 3$

$$Y_1 - 8Y_2 + 12Y_4 \leq 5$$

$$Y_1 - 9Y_2 + 7Y_3 + 4Y_4 \leq 1$$

$$Y_1, Y_2, Y_3, Y_4 \geq 0$$

$$17. \text{Max } 5Y_1 - 20Y_2 + 10Y_3$$

$$\begin{aligned} \text{subject to } Y_1 - 3Y_2 &\leq 2 \\ Y_1 + 2Y_2 + Y_3 &\leq 10 \\ Y_1 - 6Y_2 + 3Y_3 &\leq 0 \\ Y_1, Y_2, Y_3 &\geq 0 \end{aligned}$$

$$19. \text{Min } 14Y_1 + 15Y_2 - 8Y_3$$

$$\begin{aligned} \text{subject to } 2Y_1 + Y_2 - Y_3 &\geq 5 \\ Y_1 + 3Y_2 - Y_3 &\geq 3 \\ 3Y_1 + 2Y_2 - Y_3 &\geq 14 \\ Y_1, Y_2, Y_3 &\geq 0 \end{aligned}$$

8.5 Game Theory

Game theory was developed for the purpose of analyzing competitive situations involving conflicting interests. In the situations considered in game theory, there are two or more persons with different objectives, each of whose action influences, but does not completely determine, the outcome of the game; furthermore, each person is assumed to know his opponent's objectives. Game theory provides solutions to such games, assuming that each of the players wishes to maximize his minimum expected profit or, equivalently, minimize his maximum expected loss. This criterion, which is based on a somewhat conservative view of the problem, is referred to as the *minimax* or *maximin criterion*. It is the basis for the *theory of games of strategy*, originally developed by John von Neumann and Oskar Morganstern and subsequently applied in various areas of research.

Note that the theory of games assumes a particular type of situation involving maximizing the expected value of a decision made under uncertainty. Decisions made under uncertainty sometimes involve only one person making a decision; the events which, with the decision, influence the result are thought of as being controlled by some random nonrational device. Unless they are reformulated, these decision problems do not involve games of strategy, since the decision maker does not have a rational opponent. However, a minimax or game theory solution may be appropriate even for these decision problems if the decision maker wishes to protect himself against the worst possible occurrence, even though it results from a random event rather than from the decision of a rational opponent.

Most recreational games such as tick-tack-toe, checkers, backgammon, chess, poker, bridge, and other card games can be analyzed as games of strategy. As usually formulated, gambling games such as dice and roulette are not games of strategy, since a person playing one of these games is "playing against the odds," not against a rational opponent.

The applications of game theory are not limited to parlor games but include competitive situations in economics, business, warfare, and social behavior. There are several fundamental characteristics by which games are classified for solution. The most important of these are the number of persons, the nature of the payoff, and the number of available strategies. These characteristics are discussed and methods of solution of games are then considered.

Number of Persons

Games are classified according to the number of distinct sets of interests or objectives present in the game. From the game theory point of view, the number of persons in the game is not necessarily the same as the number of people playing the game—that is, if two or more players form a coalition in which they agree to pool their winnings or losses, game-theory analysis treats them as a single person.

Most of the work done thus far in game theory deals with two-person games and the discussion in the following sections is confined to this case. Analysis of games

involving more than two persons is usually very difficult. Note that many situations that are not strictly two-person games may be analyzed as though they were—for example, the interests in a game of cards can be considered as “his” and “everybody else’s.”

The Payoff

Games are also classified with respect to the nature of the payoff, that is, what happens at the end of the game. The distinction in this respect is between zero-sum games and nonzero-sum games. If the sum of the payoffs to all players of a game is zero, counting winnings as positive and losses as negative, then the game is *zero-sum*; otherwise, it is *nonzero-sum*. In a zero-sum game, anything won by one player is lost by another player. The importance of this distinction lies in the fact that a zero-sum game is a closed system and a nonzero-sum game is not. Almost all parlor games are zero-sum, and many other situations can be analyzed as zero-sum games. Most of the work in game theory has concerned zero-sum games and the discussion in the following sections is confined to such games. Note that a nonzero-sum game may be made zero-sum by adding a fictitious player, say Nature, but this necessitates a more difficult analysis, especially if the original game was a two-person game.

Strategies

In game theory, a strategy for a particular player is a plan which specifies his action for every possible action of his opponent, that is, a strategy is a complete plan for playing the game, without any connotation of skillfulness on the part of the player. In a game completely amenable to analysis, it is possible, at least conceptually, to foresee all eventualities and thus to catalogue all possible strategies. Games are classified according to the number of strategies available to each player: if player 1 has m possible strategies and player 2 has n possible strategies and they are the only players, then the game is $m \times n$, that is, m by n . The important distinction in classifying games on the basis of strategies is between finite games and infinite games. If the greatest number of strategies available to any player is finite, then the game is *finite*; if at least one player has an infinite number of available strategies, then the game is *infinite*. The theory of infinite games is very difficult and is not discussed in the following sections. For the analysis of finite games, it is convenient to distinguish three cases: those in which the player having the least number of strategies has two, three, or more than three.

In summary, the discussion in following sections will generally concern finite, zero-sum, two-person games.

The Game Matrix

A problem is usually set up for game-theory analysis in the form of a game matrix. A *game matrix* or *payoff matrix* is a rectangular array of the payoffs, where the rows represent the strategies of one player and the columns represent the strategies of the other player; thus an $m \times n$ game is represented by an $m \times n$ game matrix. It is conventional to write the payoffs from the point of view of the player whose strategies are associated with the rows of the matrix; the payoffs for the other player, in a zero-sum game, are then given by the negative of this matrix.

Game theory thus assumes that the strategies available to each player can be enumerated and that the corresponding payoffs can be expressed in meaningful, although not necessarily monetary, units. This information is sufficient for solution

of the game—that is, for determining which choice of strategies each player should make—assuming that each player wishes to maximize his minimum expected profit or minimize his maximum expected loss. The minimax theorem, the key result of the theory of games, states that such a minimax solution exists for every finite, zero-sum, two-person game. Note that minimax is not the only possible criterion for solving a game matrix and that its use leads to a conservative theory, since the opponent is assumed to be skillful and to use his best strategy.

The *value* of a game is the expected or average payoff per play over a long series of plays, assuming that both players consistently use their optimum strategies. It is conventional to regard value from the point of view of the player whose strategies correspond to the rows of the payoff matrix. A game is said to be *fair* if its value is zero; in a fair game neither player has an advantage. In a game that is not fair, one player will win from the other in the long run if both play their optimum strategies; if the value of the game is positive, the row player has the advantage; if the value of the game is negative, the column player has the advantage.

Saddle Points

If a game matrix contains an entry that is simultaneously a maximum of row minima and a minimum of column maxima, this minimax entry is said to be a *saddle point* of the game and the game is said to be *strictly determined*. In this case, according to the criterion of game theory, the *optimum strategies* for the respective players are represented by the row and column whose intersection is the saddle point. The value of a strictly determined game is the value of its saddle point.

The first step in the solution of a matrix game is to check for the existence of a saddle point; if one is found, the game is solved; if not, further analysis is necessary. Checking for a saddle point is usually done by writing the row minimum beside each row and the column maximum at the bottom of each column and then determining the maximum of the minima and the minimum of the maxima. A saddle point can also be determined by checking for an entry which is simultaneously the minimum of the row in which it occurs and the maximum of the column in which it occurs (see Table 8.13).

Table 8.13

		Player Y					
		1	2	...	n		
Player X	1	a_{11}	a_{12}	...	a_{1n}	\min_j	a_{1j}
	2	a_{21}	a_{22}	...	a_{2n}	\min_j	a_{2j}
	\vdots	\vdots	\vdots	a_{ij}	\vdots	\vdots	$\max_i \min_j a_{ij}$
	m	a_{m1}	a_{m2}	...	a_{mn}	\min_j	a_{mj}
		$\max_i a_{i1}$	$\max_i a_{i2}$...	$\max_i a_{in}$		
		$\min_j \max_i a_{ij}$					

EXAMPLE

Check the following games for saddle points.

(a)

12	2	25	-10	-10
16	3	4	10	3
-2	-1	26	0	-2
14	-4	8	6	-4
16	3	26	10	

max min = 3

min max = 3

There is a saddle point at the intersection of the second row (player X 's optimum strategy) and the second column (player Y 's optimum strategy); the value of the game is 3.

(b)

-15	22	10	8	6	-14	-8	-15
-3	4	-6	0	-4	22	-10	-10
-2	3	4	10	-1	0	-6	-6
-2	22	10	10	6	22	-6	

max min = -6

min max = -6

There is a saddle point at the intersection of the third row (player X 's optimum strategy) and the seventh column (player Y 's optimum strategy); the value of the game is -6.

(c)

-3	2	4	-3
6	1	3	1
3	10	12	3
5	0	-2	-2
0	-4	6	-4
6	10	12	

max min = 3

min max = 6

There is no saddle point.

(d)

2	0	6	-4	8	10	25	-10	12	14	-10
-14	-2	0	14	9	12	15	0	-10	-3	-14
2	0	6	14	9	12	25	0	12	14	

max min = -10

min max = 0

There is no saddle point.

SOLUTION OF TWO-PERSON, TWO-STRATEGY GAMES

The basic concepts of game-theory analysis discussed in previous sections are illustrated in the following sections for two-person, two-strategy games, that is, for games in which at least one of the two players has only two strategies. By convention, the strategies for player X are listed and indexed in a column along the left edge of the game matrix and the strategies for player Y are listed and indexed in a row along the top edge. The payoffs are to player X , that is, a positive number indicates a payoff from player Y to player X and a negative number indicates a payoff from player X to player Y .

2×2 Games

The most easily analyzed two-strategy games are 2×2 , that is, games in which each player has only two possible strategies. Solution of 2×2 subgames is also frequently necessary as a step in the solution of larger ($2 \times n$ or $m \times 2$) two-strategy games.

The existence or nonexistence of a saddle point is easily determined for a 2×2 game, either directly by the method above or using the following theorem.

THEOREM: The 2×2 matrix game

$$G = \begin{array}{|c|c|} \hline a & b \\ \hline c & d \\ \hline \end{array}$$

is nonstrictly determined if and only if one of the following two conditions is satisfied:

1. $a < b$, $a < c$, $d < b$, and $d < c$.
2. $a > b$, $a > c$, $d > b$, and $d > c$.

That is, a 2×2 matrix game is nonstrictly determined if and only if the two entries on one diagonal of the matrix are each greater than each of the two entries on the other diagonal.

EXAMPLE

The game

		Player Y	
		1	2
Player X	1	①	1
	2	-3	10

is strictly determined and fair. Player X's optimum strategy is 1; player Y's optimum strategy is 1.

EXAMPLE

The game

		Player Y	
		1	2
Player X	1	5	②
	2	-7	-4

is strictly determined, but not fair (its value is 2). Player X's optimum strategy is 1; player Y's optimum strategy is 2.

EXAMPLE

The game

		Player Y	
		1	2
Player X	1	0	1
	2	2	0

is not strictly determined.

In a nonstrictly determined game, there is no clearly optimum strategy for either player to use consistently and, furthermore, consistent use of any particular strategy by either player can be capitalized on by the other player. There is thus an important difference between strictly determined and nonstrictly determined games: In a strictly determined game there is an optimum strategy for each player and no "security measures" are necessary; in a nonstrictly determined game, optimal play involves preventing the opponent from knowing what strategy is to be used on a given play. This is accomplished by selecting the strategy to be used for each play at random, according to probabilities that can be computed from the game matrix. Such a strategy, which consists of a probability mixture of more than one (pure) strategy, is called a *mixed strategy*.

The solution of a 2×2 nonstrictly determined game consists of a pair of probabilities p_1 and $p_2 = 1 - p_1$ with which player X selects at random his strategies 1 and 2, respectively, and a pair of probabilities q_1 and $q_2 = 1 - q_1$ with which player Y selects at random his strategies 1 and 2, respectively. These probabilities provide the optimum mixed strategies, that is, the strategies by which each player, respectively, can maximize his minimum expected profits or minimize his maximum expected losses against optimal play by the other player.

The value of p_1 can be obtained by equating the two possible expected payoffs for player X and solving for p_1 , since for any other p_1 one of the two expected payoffs is less than the other and thus the minimax criterion is violated. Similarly, q_1 can be obtained by equating the two possible expected payoffs for player Y and solving.

If

$$G = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

represents the game matrix, then the expected payoff to player X if player Y uses strategy 1 is

$$ap_1 + c(1 - p_1)$$

and if player Y uses strategy 2 is

$$bp_1 + d(1 - p_1)$$

Thus, equating the expected payoffs,

$$ap_1 + c(1 - p_1) = bp_1 + d(1 - p_1)$$

$$p_1(a - b - c + d) = d - c$$

$$p_1 = \frac{d - c}{a - b - c + d}$$

$$p_2 = 1 - p_1 = \frac{a - b}{a - b - c + d}$$

Similarly, equating the (negative) expected payoffs to player Y,

$$aq_1 + b(1 - q_1) = cq_1 + d(1 - q_1)$$

$$a_1(a - b - c + d) = d - b$$

$$q_1 = \frac{d - b}{a - b - c + d}$$

$$q_2 = 1 - q_1 = \frac{a - c}{a - b - c + d}$$

The value of a game has the same meaning for strictly and nonstrictly determined games: The value of a game is the payoff that a player can expect to obtain per play; on the average, a player cannot win more than the value of a game unless his opponent plays poorly, nor can he win less than the value of a game unless he plays poorly.

The value of the game

$$G = \begin{array}{|c|c|} \hline a & b \\ \hline c & d \\ \hline \end{array}$$

(to player X) is

$$v = ap_1 + c(1 - p_1) = bp_1 + d(1 - p_1)$$

$$= -[aq_1 + b(1 - q_1)] = -[cq_1 + d(1 - q_1)]$$

$$= \frac{ad - bc}{a - b - c + d}$$

The negative of this quantity is the payoff to player Y.

EXAMPLE

The game $\begin{array}{|c|c|} \hline 2 & 0 \\ \hline 0 & 2 \\ \hline \end{array}$ is nonstrictly determined. The optimal mixed strategies are

$$p_1 = \frac{2}{4} = \frac{1}{2}$$

$$p_2 = \frac{2}{4} = \frac{1}{2}$$

$$q_1 = \frac{2}{4} = \frac{1}{2}$$

$$q_2 = \frac{2}{4} = \frac{1}{2}$$

and the value is $\frac{4}{4} = 1$; the game is biased in favor of player X.

EXAMPLE

The game $\begin{array}{|c|c|} \hline -1 & 0 \\ \hline 0 & -2 \\ \hline \end{array}$ is nonstrictly determined. The optimal mixed strategies are

$$p_1 = \frac{-2}{-3} = \frac{2}{3}$$

$$p_2 = \frac{-1}{-3} = \frac{1}{3}$$

$$q_1 = \frac{-2}{-3} = \frac{2}{3}$$

$$q_2 = \frac{-1}{-3} = \frac{1}{3}$$

and the value is $v = \frac{-2}{-3} = -\frac{2}{3}$; the game is biased in favor of player Y.

EXAMPLE

The game $\begin{bmatrix} 7 & -6 \\ 5 & 8 \end{bmatrix}$ is nonstrictly determined. The optimal mixed strategies are

$$p_1 = \frac{3}{16}$$

$$p_2 = \frac{13}{16}$$

$$q_1 = \frac{14}{16} = \frac{7}{8}$$

$$q_2 = \frac{2}{16} = \frac{1}{8}$$

and the value is $v = \frac{86}{16} = 5\frac{3}{8}$; the game is biased in favor of player X .

EXAMPLE

The game $\begin{bmatrix} 10 & -30 \\ -10 & 20 \end{bmatrix}$ is nonstrictly determined. The optimal mixed strategies are

$$p_1 = \frac{30}{70} = \frac{3}{7}$$

$$p_2 = \frac{40}{70} = \frac{4}{7}$$

$$q_1 = \frac{50}{70} = \frac{5}{7}$$

$$q_2 = \frac{20}{70} = \frac{2}{7}$$

and the value is $v = -\frac{100}{70} = -1\frac{3}{7}$; the game is biased in favor of player Y .

The play of a game is not affected by adding a constant to all payoffs or by multiplying all payoffs by a positive constant. The value of the game is affected by the same transformation as that applied to the payoffs of the game matrix.

EXAMPLE

For each of the games

$$G_1 = \begin{bmatrix} 8 & 1 \\ 4 & 6 \end{bmatrix}$$

$$G_2 = \begin{bmatrix} 11 & 4 \\ 7 & 9 \end{bmatrix}$$

$$G_3 = \begin{bmatrix} 16 & 2 \\ 8 & 12 \end{bmatrix}$$

$p_1 = \frac{2}{9}$, $p_2 = \frac{7}{9}$, $q_1 = \frac{5}{9}$, $q_2 = \frac{4}{9}$. Note that $G_2 = G_1 + 3$; the value of G_1 is $4\frac{8}{9}$ and the value of G_2 is $7\frac{8}{9}$. $G_3 = 2G_1$; the value of G_1 is $4\frac{8}{9}$ and the value of G_3 is $9\frac{7}{9}$.

Solution of 2×2 Games by Matrix Algebra

The optimum strategies and the value of a nonstrictly determined 2×2 game can be obtained using matrix algebra as follows: If the payoff matrix is represented by

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

then X 's optimal strategies are given by

$$[p_1, p_2] = \frac{[1, 1][\text{adj } A]}{[1, 1][\text{adj } A] \begin{bmatrix} 1 \\ 1 \end{bmatrix}}$$

Y 's optimal strategies are given by

$$[q_1, q_2] = \frac{[1, 1][\text{adj } A]'}{[1, 1][\text{adj } A] \begin{bmatrix} 1 \\ 1 \end{bmatrix}}$$

and the value of the game is given by

$$v = \frac{|A|}{[1, 1][\text{adj } A] \begin{bmatrix} 1 \\ 1 \end{bmatrix}}$$

Alternatively, the value of the game is given by

$$v = [p_1, p_2] \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix}$$

The examples on pages 645 and 646 can be used to illustrate this method of solution.

EXAMPLE

The game $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ is nonstrictly determined.

$$\begin{aligned} [p_1, p_2] &= \frac{[1, 1][\text{adj } A]}{[1, 1][\text{adj } A] \begin{bmatrix} 1 \\ 1 \end{bmatrix}} \\ &= \frac{[1, 1] \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}}{[1, 1] \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}} = \frac{[2, 2]}{4} = \left[\frac{1}{2}, \frac{1}{2} \right] \end{aligned}$$

$$\begin{aligned} [q_1, q_2] &= \frac{[1, 1][\text{adj } A]'}{[1, 1][\text{adj } A] \begin{bmatrix} 1 \\ 1 \end{bmatrix}} \\ &= \frac{[1, 1] \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}}{[1, 1] \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}} = \frac{[2, 2]}{4} = \left[\frac{1}{2}, \frac{1}{2} \right] \end{aligned}$$

$$\begin{aligned} v &= \frac{|A|}{[1, 1][\text{adj } A] \begin{bmatrix} 1 \\ 1 \end{bmatrix}} \\ &= \frac{4}{4} = 1 \end{aligned}$$

Alternatively,

$$\begin{aligned}
 v &= [p_1, p_2] \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \\
 &= \left[\frac{1}{2}, \frac{1}{2}\right] \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \\
 &= [1, 1] \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \\
 &= 1
 \end{aligned}$$

(as obtained above using algebraic methods).

EXAMPLE

The game

-1	0
0	-2

 is nonstrictly determined.

$$\begin{aligned}
 [p_1, p_2] &= \frac{[1, 1][\text{adj } A]}{[1, 1][\text{adj } A] \begin{bmatrix} 1 \\ 1 \end{bmatrix}} \\
 &= \frac{[1, 1] \begin{bmatrix} -2 & 0 \\ 0 & -1 \end{bmatrix}}{[1, 1] \begin{bmatrix} -2 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}} = \frac{[-2, -1]}{-3} = \left[\frac{2}{3}, \frac{1}{3}\right]
 \end{aligned}$$

$$\begin{aligned}
 [q_1, q_2] &= \frac{[1, 1][\text{adj } A]'}{[1, 1][\text{adj } A] \begin{bmatrix} 1 \\ 1 \end{bmatrix}} \\
 &= \frac{[1, 1] \begin{bmatrix} -2 & 0 \\ 0 & -1 \end{bmatrix}'}{[1, 1] \begin{bmatrix} -2 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}} = \frac{[-2, -1]}{-3} = \left[\frac{2}{3}, \frac{1}{3}\right]
 \end{aligned}$$

$$\begin{aligned}
 v &= \frac{|A|}{[1, 1][\text{adj } A] \begin{bmatrix} 1 \\ 1 \end{bmatrix}} \\
 &= \frac{2}{-3} = -\frac{2}{3}
 \end{aligned}$$

Alternatively,

$$\begin{aligned}
 v &= [p_1, p_2] \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \\
 &= \left[\frac{2}{3}, \frac{1}{3}\right] \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix} \\
 &= \left[-\frac{2}{3}, -\frac{2}{3}\right] \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix} \\
 &= -\frac{2}{3}
 \end{aligned}$$

(as obtained above using algebraic methods).

EXAMPLE

The game $\begin{bmatrix} 7 & -6 \\ 5 & 8 \end{bmatrix}$ is nonstrictly determined.

$$\begin{aligned} [p_1, p_2] &= \frac{[1, 1][\text{adj } A]}{[1, 1][\text{adj } A] \begin{bmatrix} 1 \\ 1 \end{bmatrix}} \\ &= \frac{[1, 1] \begin{bmatrix} 8 & 6 \\ -5 & 7 \end{bmatrix}}{[1, 1] \begin{bmatrix} 8 & 6 \\ -5 & 7 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}} = \frac{[3, 13]}{16} = \left[\frac{3}{16}, \frac{13}{16} \right] \end{aligned}$$

$$\begin{aligned} [q_1, q_2] &= \frac{[1, 1][\text{adj } A]'}{[1, 1][\text{adj } A] \begin{bmatrix} 1 \\ 1 \end{bmatrix}} \\ &= \frac{[1, 1] \begin{bmatrix} 8 & -5 \\ 6 & 7 \end{bmatrix}}{[1, 1] \begin{bmatrix} 8 & 6 \\ -5 & 7 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}} = \frac{[14, 2]}{16} = \left[\frac{7}{8}, \frac{1}{8} \right] \end{aligned}$$

$$\begin{aligned} v &= \frac{|A|}{[1, 1][\text{adj } A] \begin{bmatrix} 1 \\ 1 \end{bmatrix}} \\ &= \frac{86}{16} = \frac{43}{8} \end{aligned}$$

Alternatively,

$$\begin{aligned} v &= [p_1, p_2] \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \\ &= \left[\frac{3}{16}, \frac{13}{16} \right] \begin{bmatrix} 7 & -6 \\ 5 & 8 \end{bmatrix} \begin{bmatrix} \frac{7}{8} \\ \frac{1}{8} \end{bmatrix} \\ &= \left[\frac{43}{8}, \frac{43}{8} \right] \begin{bmatrix} \frac{7}{8} \\ \frac{1}{8} \end{bmatrix} \\ &= \frac{43}{8} \end{aligned}$$

(as obtained above using algebraic methods).

EXAMPLE

The game $\begin{bmatrix} 10 & -30 \\ -10 & 20 \end{bmatrix}$ is nonstrictly determined.

$$\begin{aligned} [p_1, p_2] &= \frac{[1, 1][\text{adj } A]}{[1, 1][\text{adj } A] \begin{bmatrix} 1 \\ 1 \end{bmatrix}} \\ &= \frac{[1, 1] \begin{bmatrix} 20 & 30 \\ 10 & 10 \end{bmatrix}}{[1, 1] \begin{bmatrix} 20 & 30 \\ 10 & 10 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}} = \frac{[30, 40]}{70} = \left[\frac{3}{7}, \frac{4}{7} \right] \end{aligned}$$

$$\begin{aligned}
[q_1, q_2] &= \frac{[1, 1][\text{adj } A]'}{[1, 1][\text{adj } A] \begin{bmatrix} 1 \\ 1 \end{bmatrix}} \\
&= \frac{[1, 1] \begin{bmatrix} 20 & 10 \\ 30 & 10 \end{bmatrix}}{[1, 1] \begin{bmatrix} 20 & 30 \\ 10 & 10 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}} = \frac{[50, 20]}{70} = \left[\frac{5}{7}, \frac{2}{7} \right] \\
v &= \frac{|A|}{[1, 1][\text{adj } A] \begin{bmatrix} 1 \\ 1 \end{bmatrix}} \\
&= \frac{-100}{70} = -\frac{10}{7}
\end{aligned}$$

Alternatively,

$$\begin{aligned}
v &= [p_1, p_2] \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \\
&= \left[\frac{3}{7}, \frac{4}{7} \right] \begin{bmatrix} 10 & -30 \\ -10 & 20 \end{bmatrix} \begin{bmatrix} \frac{5}{7} \\ \frac{2}{7} \end{bmatrix} \\
&= \left[-\frac{10}{7}, -\frac{10}{7} \right] \begin{bmatrix} \frac{5}{7} \\ \frac{2}{7} \end{bmatrix} \\
&= -\frac{10}{7}
\end{aligned}$$

(as obtained above using algebraic methods).

$2 \times n$ Games and $m \times 2$ Games

In $2 \times n$ games and $m \times 2$ games, one player has two strategies and the other player has more than two strategies; the solution of a $2 \times n$ game or an $m \times 2$ game can be reduced to the solution of a 2×2 subgame.

As in the solution of 2×2 games, the first step is to check for a saddle point if one exists; the optimum (pure) strategies and the value of the game are thus determined.

EXAMPLE

The game

4	4
5	3
6	⑤
1	3
5	4

has a saddle point at the intersection of the third row (player X 's optimal strategy) and the second column (player Y 's optimal strategy); the value of the game is 5.

EXAMPLE

The game

1	7	①	3
4	8	-1	6

has a saddle point at the intersection of the first row (player X 's optimal strategy) and the third column (player Y 's optimal strategy); the value of the game is zero.

Dominance

In an $m \times n$ matrix game, row i is said to *majorize* or to *dominate* row h if every entry in row i is as large or larger than the corresponding entry in row h . Similarly, column j is said to *minorize* or to *dominate* column k if every entry in column j is as small as or smaller than the corresponding entry in column k . Note that any dominated (majorized) row or dominated (minorized) column can be omitted from the matrix game without affecting its solution, since such strategies are clearly not optimal. If a $2 \times n$ or an $m \times 2$ game does not have a saddle point, that is, is not strictly determined, all majorized rows and minorized columns should be eliminated as the next step in the solution.

The solution of a $2 \times n$ game consists of probabilities p_1 and $p_2 = 1 - p_1$ with which player X selects, at random, his strategies 1 and 2, respectively, and probabilities q_1, q_2, \dots, q_n , where $\sum_{i=1}^n q_i = 1$, with which player Y selects, at random, his strategies 1, 2, \dots , n , respectively. Similarly, the solution of an $m \times 2$ game consists of probabilities p_1, p_2, \dots, p_m , where $\sum_{i=1}^m p_i = 1$, for player X and probabilities q_1 and $q_2 = 1 - q_1$ for player Y .

EXAMPLE

The game

2	5
4	3
3	6
5	4
4	4

has no saddle point. However, row 3 dominates row 1 and row 4 dominates rows 2 and 5. Thus the game is reduced, for calculation, to the subgame

3	6
5	4

EXAMPLE

The game

-6	-1	1	4	7	4	3
7	-2	6	3	-2	-5	7

has no saddle point. However, columns 3, 4, 5, and 7 dominate column 2. Thus the game is reduced, for calculation, to the subgame

-6	-1	4
7	-2	-5

After dominance has been used to reduce a $2 \times n$ or an $m \times 2$ game for calculation, all possible 2×2 games derived from the matrix of this reduced game can be solved. The value of the original game is the value of one of these derived 2×2

games and the optimal strategies of the original game are also those of that derived 2×2 game, extended for one player by the addition of zeros. Which of the derived games provides the solution of the original game can be determined either by trial and error or graphically.

The trial-and-error procedure consists of solving the derived 2×2 games until one is found for which the two-strategy player does at least as well (usually better) against all his opponent's other strategies as he does against the pair appearing in the 2×2 subgame. When such a game is determined, its solution provides the solution of the original game.

EXAMPLE

In the first example above, only one 2×2 subgame remained

3	6
5	4

Its solution is $p_1 = \frac{1}{4}$, $p_2 = \frac{3}{4}$; $q_1 = \frac{1}{2}$, $q_2 = \frac{1}{2}$; and the solution of the 5×2 game is thus $p_1 = 0$, $p_2 = 0$, $p_3 = \frac{1}{4}$, $p_4 = \frac{3}{4}$, $p_5 = 0$; $q_1 = \frac{1}{2}$, $q_2 = \frac{1}{2}$. For both games, $v = \frac{9}{2}$.

EXAMPLE

In the second example above, a 2×3 subgame remained

-6	-1	4
7	-2	-5

and thus there are three 2×2 games for possible solution. The first of these 2×2 games,

-6	-1
7	-2

has the solution $p_1 = \frac{9}{14}$, $p_2 = \frac{5}{14}$; $q_1 = \frac{1}{14}$, $q_2 = \frac{13}{14}$; its value is $-\frac{19}{14}$.

Against the other remaining strategy of player Y, column 3 of the reduced game, $p_1 = \frac{9}{14}$, $p_2 = \frac{5}{14}$ has the value

$$\frac{9}{14}(4) + \frac{5}{14}(-5) = \frac{11}{14}$$

which is greater than $-\frac{19}{14}$, so the solution of this 2×2 game extended is the solution of the original game:

$$p_1 = \frac{9}{14} \quad p_2 = \frac{5}{14}$$

$$q_1 = \frac{1}{14} \quad q_2 = \frac{13}{14} \quad q_3 = q_4 = q_5 = q_6 = q_7 = 0$$

NOTE: The 2×2 game

-1	4
-2	-5

has a saddle point; thus $p_1 = 1$, $p_2 = 0$, and the value of the game is -1 . Against the other remaining strategy of player Y, column 1 of the reduced game, $p_1 = 1$, $p_2 = 0$ has the value -6 , which is *not* greater than -1 , so the solution of this game is *not* the solution of the original game. The 2×2 game

-6	4
7	-5

has the solution $p_1 = \frac{6}{11}$, $p_2 = \frac{5}{11}$, and the value of the game is $-\frac{1}{11}$. Against the other remaining strategy of player Y, column 2 of the reduced game, $p_1 = \frac{6}{11}$, $p_2 = \frac{5}{11}$ has value $-\frac{6}{11}$, which is *not* greater than $-\frac{1}{11}$, so the solution of this game is *not* the solution of the original game.

Graphical Solution

Graphically, the 2×2 game whose solution is the solution of a $2 \times n$ game is determined as follows. Plot the pairs of payoffs of the n strategies of player Y on two vertical axes and connect the pairs of points by straight lines; locate the highest point on the line segments that form the lower boundary of the graph. The lines that intersect at this point identify the strategies player Y should use in his optimum strategy. Note that it may be possible to reduce a game by dominance before doing graphical analysis.

EXAMPLE

In the game

2	-2	3	7	6
6	5	1	4	0

column 5 dominates column 4 and column 2 dominates column 1. The 2×3 game to be solved is

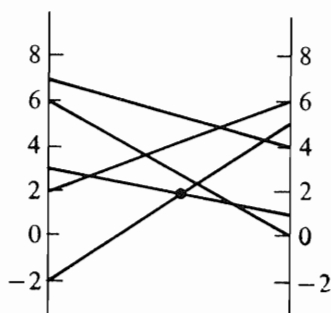
-2	3	6
5	1	0

which can be solved by solving the 2×2 game

-2	3
5	1

(See Figure 8.5.)

Figure 8.5



The solution of this 2×2 subgame is $p_1 = \frac{4}{9}$, $p_2 = \frac{5}{9}$; $q_1 = \frac{2}{9}$, $q_2 = \frac{7}{9}$; its value is $\frac{17}{9}$. Thus the solution of the original game is $p_1 = \frac{4}{9}$, $p_2 = \frac{5}{9}$; $q_1 = 0$, $q_2 = \frac{2}{9}$, $q_3 = \frac{7}{9}$, $q_4 = 0$, $q_5 = 0$; its value is $\frac{17}{9}$.

Similarly, for an $m \times 2$ game, the payoffs of the m strategies of player X are plotted and the lines that intersect at the lowest point on the line segments that form the upper boundary of the figure identify the strategies player X should use in his optimum strategy.

EXAMPLE

In the game

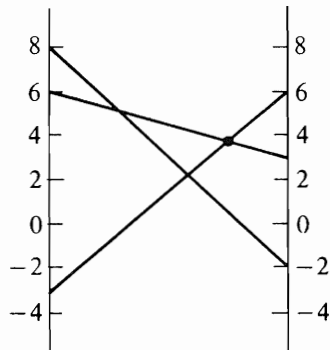
-3	6
6	3
8	-2

the 2×2 game to be solved is (see Figure 8.6)

-3	6
6	3

The solution of this 2×2 subgame is $p_1 = \frac{1}{4}$, $p_2 = \frac{3}{4}$; $q_1 = \frac{1}{4}$, $q_2 = \frac{3}{4}$; its value is $\frac{15}{4}$. Thus the solution of the original game is $p_1 = \frac{1}{4}$, $p_2 = \frac{3}{4}$, $p_3 = 0$; $q_1 = \frac{1}{4}$, $q_2 = \frac{3}{4}$; its value is $\frac{15}{4}$.

Figure 8.6



EXAMPLE

An automobile manufacturer has five proposed designs for next year's new cars. Which of these is likely to sell best depends largely on whether his competitor's standard model is excellent, good, fair, or poor. If the model is excellent, his net profit (millions of dollars) will be 100, 150, 50, 125, and 90, respectively; if the model is good, his profit will be 80, 55, 55, 60, and 70, respectively; if the model is fair, his profit will be 150, 100, 100, 100, and 125, respectively; if the model is poor, his profit will be 50, 80, 25, 80, and 75, respectively. What design should he choose in order to maximize his minimum expected profit?

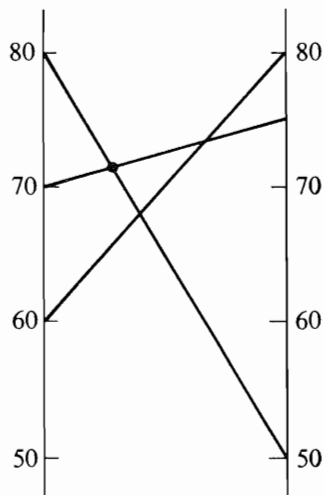
		Competitor's model			
		Excellent	Good	Fair	Poor
Design	1	100	80	150	50
	2	150	55	100	80
	3	50	55	100	25
	4	125	60	100	80
	5	90	70	125	75

The 2×2 game to be solved is (see Figure 8.7)

80	50
70	75

The solution of this 2×2 subgame is $p_1 = \frac{1}{7}$, $p_2 = \frac{6}{7}$ and $q_1 = \frac{5}{7}$, $q_2 = \frac{2}{7}$. Thus the manufacturer should produce model 1 with probability $\frac{1}{7}$ and model 5 with probability $\frac{6}{7}$; his competitor

Figure 8.7



should produce a good model with probability $\frac{5}{7}$ and a poor model with probability $\frac{2}{7}$; the value is $\frac{190}{7}$. Note that the payoffs in this problem are all positive and the competitor must therefore have some external reason (such as trying to enter the market) for manufacturing at an inevitable loss. This is assuming that the matrix literally represents net profits and not some type of additional revenue.

EXAMPLE

The Defense Department plans to award a contract for a new missile range at one of two locations, *A* or *B*. A real estate speculator intends to invest \$5000 in land—all at location *A*, all at location *B*, or half at each location. If he buys at location *A*, the land will be worth \$10,000 if the missile range is built there, but \$3000 if it is built at location *B*. If he buys at location *B*, the land will be worth \$4000 if the missile range is built at location *A* and \$8000 if it is built at location *B*. If he buys at both locations, the land will be worth \$6000 if the missile range is built at location *A* and \$5000 if it is built at location *B*. In order to maximize his minimum expected profits, what should the speculator do?

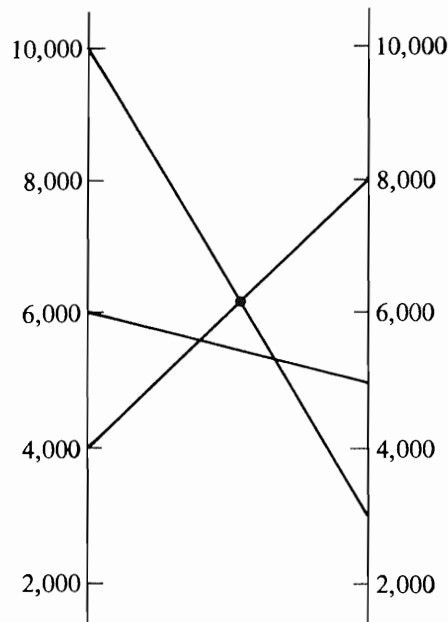
		Missile range	
		<i>A</i>	<i>B</i>
Investment	<i>A</i>	10,000	3,000
	<i>B</i>	4,000	8,000
	<i>A</i> and <i>B</i>	6,000	5,000

The 2×2 subgame to be solved is (see Figure 8.8)

10,000	3,000
4,000	8,000

The solution of this 2×2 subgame is $p_1 = \frac{4}{11}$, $p_2 = \frac{7}{11}$. Thus the speculator should invest his money in property *A* with probability $\frac{4}{11}$ and property *B* with probability $\frac{7}{11}$; the value is \$6181.82. If the Defense Department is considered to be somehow plotting against the speculator, it should choose location *A* with probability $\frac{5}{11}$ and location *B* with probability $\frac{6}{11}$.

Figure 8.8



Alternatively, the speculator can be considered to be choosing in order to protect himself against the worst possible occurrence, whether that occurrence is in fact random or is the result of Defense Department strategy against him.

SOLUTION OF LARGER GAMES

For 3×3 and larger games, if there is no saddle point and if the original game cannot be reduced to a smaller game by dominance, linear programming provides the most efficient method of solution. Linear programming is concerned with the problem of maximizing or minimizing a linear function whose variables are restricted to values satisfying a system of linear constraints. A matrix game can be expressed as a problem of this type, since each player is concerned with maximizing the value of the game to himself but is subject to the constraints imposed by the payoff matrix.

If, as above, the payoff matrix is represented by

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

and Y 's optimum strategies are represented by (q_1, q_2, \dots, q_n) , the inequalities which express the expectations of Y are given by

$$a_{11}q_1 + a_{12}q_2 + \cdots + a_{1n}q_n \leq v$$

$$a_{21}q_1 + a_{22}q_2 + \cdots + a_{2n}q_n \leq v$$

$$\vdots$$

$$a_{m1}q_1 + a_{m2}q_2 + \cdots + a_{mn}q_n \leq v$$

$$q_1 + q_2 + \cdots + q_n = 1$$

Substituting $\bar{q}_i = q_i/v$,

$$a_{11}\bar{q}_1 + a_{12}\bar{q}_2 + \cdots + a_{1n}\bar{q}_n \leq 1$$

$$a_{21}\bar{q}_1 + a_{22}\bar{q}_2 + \cdots + a_{2n}\bar{q}_n \leq 1$$

$$\vdots$$

$$a_{m1}\bar{q}_1 + a_{m2}\bar{q}_2 + \cdots + a_{mn}\bar{q}_n \leq 1$$

$$\bar{q}_1 + \bar{q}_2 + \cdots + \bar{q}_n = \frac{1}{v}$$

Player Y's objective is to minimize v or, equivalently, maximize $1/v$. Thus the solution of the matrix game can be written as a linear programming problem as follows:

$$\text{Maximize } \bar{q}_1 + \bar{q}_2 + \cdots + \bar{q}_n$$

subject to

$$a_{11}\bar{q}_1 + a_{12}\bar{q}_2 + \cdots + a_{1n}\bar{q}_n \leq 1$$

$$a_{21}\bar{q}_1 + a_{22}\bar{q}_2 + \cdots + a_{2n}\bar{q}_n \leq 1$$

$$\vdots$$

$$a_{m1}\bar{q}_1 + a_{m2}\bar{q}_2 + \cdots + a_{mn}\bar{q}_n \leq 1$$

$$\bar{q}_1 + \bar{q}_2 + \cdots + \bar{q}_n = \frac{1}{v}$$

or, more concisely,

$$\text{Maximize } \sum_{i=1}^n \bar{q}_i$$

subject to

$$\sum_{j=1}^n a_{ij}\bar{q}_j \leq 1 \quad \text{for } i = 1, 2, \dots, m$$

$$\sum_{j=1}^n \bar{q}_j = \frac{1}{v}$$

PROBLEMS

Find the optimal strategies for the players of each of the following matrix games using the minimax criterion.

1.

8	10	13	16	9
10	12	6	15	10
16	18	9	13	25
4	9	18	20	6

2.

5	8	7
-1	-3	10
2	12	-6

3.

10	8	6	2
15	12	2	4
-4	6	-3	1
12	-2	8	-6
16	13	7	12

4.

14	3	5	8	4
-3	-1	6	-2	5
2	12	10	13	16

5.

15	-20	-12	3
4	2	-10	-6
20	-18	-15	-8
-12	8	-10	6
10	9	-11	4

6.

3	10	20	18	0
-9	-8	46	10	4
2	0	17	18	0
-1	1	13	5	3

Find the optimal strategies for the players of each of the following matrix games using the minimax criterion and determine the value of each game.

7.

7	-1	7
10	-2	-5
9	5	6

8.

3	5	2	1
0	4	8	3
2	7	14	9

9.

5	4	-2	1
7	-6	3	6
12	8	10	9
6	18	-9	14

10.

3	-10	-5
2	-4	-3
4	-5	-2
-6	-4	0
10	-8	1

11.

12	-10	8	-6	7	-11
18	2	-3	-4	8	10
5	-3	14	0	-10	-12
3	12	16	1	8	2
-15	16	12	-2	9	-9

12.

10	8	11	-2
14	6	-5	5
9	7	5	-4
15	4	-3	3

13.

4	6
9	5
3	1
7	8
10	2

14.

6	-10	12	25	0	5	3	-2
2	-8	4	-6	-7	9	4	10

15.

0	2
1	3
-1	0
2	0

16.

9	-5	7	1	-3
-10	4	-8	-6	2

17. A contractor is going to build a large number of houses for a housing development. Four types of houses have been discussed: colonial, ranch, split-level, and contemporary; the development committee will choose two of these types for the contractor to construct. The contractor has the opportunity to buy materials in carload lots, thus saving considerable money, but he must order in advance of the committee's decision and can order only one type of material. If the committee chooses colonial and ranch, the contractor will make an (extra) profit (in thousands of dollars) of 125, 120, 60, and 50 if he orders colonial, ranch, split-level, and contemporary materials, respectively. If the committee chooses colonial and split-level, he will make 90, 40, 80, 75, respectively; if the committee chooses colonial and contemporary, he will make 150, 30, 75, 100, respectively; if the committee chooses ranch and split level, he will make 70, 70, 75, 65, respectively; if the committee chooses ranch and contemporary, he will make 90, 80, 80, 120, respectively; if the committee chooses split-level and contemporary, he will make 80, 40, 130, and 80, respectively. How should the contractor order to maximize his minimum expected extra profit?
18. A plant manager must set up his reactors to produce a certain type of polythene using either process 1, 2, 3, 4, or 5. Unfortunately, the chemical raw material varies in nitrogen

content and may contain 3, 4, 5, or 6% nitrogen. The nitrogen content affects the relative efficiency of the five processes. With 3% nitrogen, the five processes have an output of 50, 45, 60, 50, and 30 tons, respectively. With 4% nitrogen, the outputs are 60, 70, 75, 90, and 60 tons, respectively; with 5% nitrogen, the outputs are 30, 55, 60, 45, and 70 tons, respectively; with 6% nitrogen, the outputs are 45, 80, 80, 65, and 85 tons, respectively. Testing for nitrogen content is too expensive to be practical. What process should the plant manager use to maximize the minimum expected output?

19. A firm builds construction machinery whose performance depends on the reliability of a gasoline motor. The firm can buy an expensive motor, fully guaranteed including replacement costs, for \$500; the firm can also buy a moderately expensive motor for \$400, guaranteed for half its cost and the cost of replacement, so that a failure means the firm pays \$600 for the motor; or the firm can buy a cheaper motor for \$300; if it buys this cheaper motor it can either guarantee replacement for a total cost of \$700, or it can pay \$50 to have the motor examined before installation, so total cost, if the motor is defective, is \$650. To minimize maximum expected costs, what should the firm do?
20. Mr. Smith and Mr. Jones each have a small hothouse in which they grow plants to sell. Each of them grows either tomatoes, flowers, or strawberries in any particular year but they are very secretive about their plans. If Mr. Smith grows tomatoes, his profit is \$100 if Mr. Jones grows tomatoes, \$150 if Mr. Jones grows strawberries, and \$200 if Mr. Jones grows flowers. If Mr. Smith grows strawberries, his profit is \$180, \$125, \$200 if Mr. Jones grows tomatoes, strawberries, flowers, respectively. If Mr. Smith grows flowers, his profit is \$140, \$125, \$100 if Mr. Jones grows tomatoes, strawberries, flowers, respectively. For each to maximize his minimum expected profit, what should Mr. Smith and Mr. Jones grow?
21. Mr. Reno has decided to bet \$5 on a basketball game between Anglewood and Barleyville. Mr. Las and Mr. Vegas have each offered Mr. Reno a bet. Mr. Las wants to bet on Barleyville and is willing to pay Mr. Reno \$10 if Anglewood wins and collect \$5 if Barleyville wins. Mr. Vegas wants to bet on Anglewood and is willing to pay Mr. Reno \$4 if Barleyville wins and collect \$5 if Anglewood wins. If he wants to maximize his minimum expected winnings, what should Mr. Reno do?
22. Casey has a concession at the Yankee Stadium for the sale of sunglasses and umbrellas. He has observed that he can sell about 500 umbrellas when it rains, and about 100 when it shines; and in the latter case he can also dispose of 1000 sunglasses. Umbrellas cost \$0.50 and sell for \$1.00; glasses cost \$0.20 and sell for \$0.50. He is willing to invest \$250 in the project. Assuming that everything that isn't sold is a total loss, what should he purchase to maximize his minimum expected profit?
23. Mr. Hilton Conrad is faced with the following decision: The State Medical Society would like to use his hotel for their annual convention; they would essentially take over all facilities of the hotel (so business from other sources would be negligible) and would pay \$20,000 for this privilege. Ordinarily the hotel would do \$10,000 business during the time the convention would be in session; however, there is a chance that the World Series will be played in the city at that time and, in this event, the hotel would do \$50,000 business. Should Mr. Hilton Conrad accept or decline the Medical Society's offer in order to maximize his minimum expected business?
24. An investor has a choice of buying one of three stocks: *A*, *B*, or *C*. The outcome of his purchase depends on whether a particular company executes a merger, divests itself of a subsidiary, or maintains the status quo. In case of a merger, stock *A* results in a gain of 20, stock *B* in a loss of 25, and stock *C* in a gain of 12. In case of divesting, stock *A* results in a loss of 5, stock *B* in a loss of 10, and stock *C* in a loss of 12. If the status quo is maintained, stock *A* results in a gain of 5, stock *B* in a gain of 30, and stock *C* in a loss of 4 (amounts are in thousands of dollars). a. In order to maximize his minimum expected gains, what stock should the investor buy? b. If divesting were impossible, but the gains and losses from the other possibilities were unchanged, what stock should he buy to maximize his minimum expected gains?
25. A commuter must decide what to do about buying insurance for his car. He will definitely carry liability insurance, but the car is rather old and he is not sure whether collision insurance is worthwhile. He can either carry no collision insurance, buy a \$50 deductible

- policy for \$60, or buy a full-coverage policy for \$70. The commuter drives this car only to and from work and he thinks one of three things will happen during the year: He will have no accident, he will have one minor accident not exceeding \$50 damage, or he will have one more serious accident not exceeding \$250 damage (the car is worth only about \$400 so this is not unreasonable optimism). The commuter realizes he could possibly have more than one accident during the year, but on the basis of his past performances he is willing to assume that he won't. What should he do to minimize his maximum expected cost?
26. In the summer Mr. Smith is considering the winter coal problem. During a normal winter, it takes about 15 tons of coal to heat his shop, but he has observed extremes where as little as 10 tons or as much as 20 tons were used. The price per ton seems to fluctuate with the weather, being \$10, \$15, and \$20 a ton during mild, normal, and severe winters, respectively. He can buy now at \$10 a ton. Mr. Smith considers three possible alternatives: to buy 10, 15, or 20 tons now and the rest, if any, later. Assuming that all coal not used is a total loss (Mr. Smith plans to sell the shop in the spring), what should he do to minimize his maximum expected cost?
27. A student must decide how to study for a final exam in history. He studies differently for true-false, multiple-choice, and essay exams, and he doesn't know which type this exam will be. The student thinks that if he studies for a true-false test, he will score 85 on a true-false test, 80 on a multiple-choice test, and 75 on an essay test. If he studies for a multiple-choice test, he will score 85 on a true-false test, 90 on a multiple-choice test, and 85 on an essay test. If he studies for an essay test, he will score 80 on a true-false test, 90 on a multiple-choice test, and 90 on an essay test. To maximize his minimum expected score, for what type of test should the student study?
28. A service station must have its service and parking area plowed after every heavy snowfall. The manager can pay \$5 every time the area needs plowing; he can buy a contract for \$25 which provides plowing for up to and including six snowfalls, with each additional plowing at a cost of \$3; or he can buy a contract for \$30 which provides plowing as many times as it is needed. What should the manager do to minimize his maximum expected cost of plowing if he is willing to assume (based on past records) that there will be between 3 and 8 (inclusive) snowfalls during the winter?

Answers to Odd-Numbered Problems

1. $p_1 = p_2 = 0, p_3 = \frac{2}{3}, p_4 = \frac{1}{3}$
 $q_1 = \frac{3}{7}, q_2 = 0, q_3 = \frac{4}{7}, q_4 = q_5 = 0$
3. $p_1 = p_2 = p_3 = 0, p_4 = \frac{5}{19}, p_5 = \frac{14}{19}$
 $q_1 = q_2 = 0, q_3 = \frac{18}{19}, q_4 = \frac{1}{19}$
5. $p_1 = 0, p_2 = 1, p_3 = p_4 = p_5 = 0$
 $q_1 = q_2 = 0, q_3 = 1, q_4 = 0$
7. $p_1 = p_2 = 0, p_3 = 1$
 $q_1 = 0, q_2 = 1, q_3 = 0$
 $v = 5$
9. $p_1 = p_2 = 0, p_3 = \frac{27}{29}, p_4 = \frac{2}{29}$
 $q_1 = 0, q_2 = \frac{19}{29}, q_3 = \frac{10}{29}, q_4 = 0$
 $v = \frac{252}{29}$
11. $p_1 = p_2 = p_3 = 0, p_4 = 1, p_5 = 0$
 $q_1 = q_2 = q_3 = 0, q_4 = 1, q_5 = q_6 = 0$
 $v = 1$
13. $p_1 = 0, p_2 = \frac{1}{5}, p_3 = 0, p_4 = \frac{4}{5}, p_5 = 0$
 $q_1 = \frac{3}{5}, q_2 = \frac{2}{5}$
 $v = \frac{37}{5}$
15. $p_1 = 0, p_2 = \frac{1}{2}, p_3 = 0, p_4 = \frac{1}{2}$
 $q_1 = \frac{3}{4}, q_2 = \frac{1}{4}$
 $v = \frac{3}{2}$
17. buy for colonial with probability $\frac{3}{14}$ and for split-level with probability $\frac{11}{14}$
19. buy expensive motor

21. bet Las with probability $\frac{3}{8}$ and Vegas with probability $\frac{5}{8}$
23. accept offer of a convention
25. buy full coverage
27. study for multiple-choice test

8.6 First-Order Markov Processes

Sequences of observations or experimental outcomes are sometimes considered to be independent; that is, the probability of observing any particular outcome is assumed to be constant. The simplest generalization of this model permits the probability of the outcome for any experiment or observation to depend on the outcome of the immediately preceding observation, but not on the outcomes of other prior observations. A process or sequence of this type is said to be a *first-order Markov chain process*, *first-order Markov chain*, or *first-order Markov process*. First-order Markov processes and their steady or equilibrium states are defined and illustrated in this section.

DEFINITION OF A FIRST-ORDER MARKOV PROCESS

Suppose that the outcome of each of a sequence of experiments or observations is one of a finite number of possible outcomes a_1, a_2, \dots, a_r . The probability of outcome a_j for any given experiment or observation depends on at most the outcome of the immediately preceding observation. These probabilities are denoted by p_{ij} , $i = 1, 2, \dots, r$ and $j = 1, 2, \dots, r$, where p_{ij} represents the probability of outcome a_j for any particular observation given that outcome a_i occurred for the immediately preceding observation. The outcomes a_1, a_2, \dots, a_r are called *states* and the p_{ij} are called *transition probabilities* of a first-order Markov chain. If it is assumed that the process begins in some particular state, the probabilities of various sequences of observations can be calculated. Thus a first-order Markov chain is specified by defining its possible states, specifying the initial probability distribution for these states, and specifying the transition matrix.

The transition probabilities can be summarized in a square matrix. For a Markov process with states a_1, a_2, \dots, a_r , the matrix of transition probabilities is

$$\mathbf{P} = \{p_{ij}\} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1r} \\ p_{21} & p_{22} & \cdots & p_{2r} \\ \vdots & \vdots & & \vdots \\ p_{r1} & p_{r2} & \cdots & p_{rr} \end{bmatrix}$$

Note that the sum of the elements in each row of the matrix \mathbf{P} is 1, since the elements of the i th row represent the probabilities for all possible transitions when the process is in state a_i . That is,

$$\sum_{j=1}^r p_{ij} = 1 \quad \text{for } i = 1, 2, \dots, r$$

Thus if the probability distribution of the states on trial n is $[p_1, p_2, \dots, p_r]$, the probability distribution of the states on trial $n + 1$ is

$$[p_1, p_2, \dots, p_r] \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1r} \\ p_{21} & p_{22} & \cdots & p_{2r} \\ \vdots & \vdots & & \vdots \\ p_{r1} & p_{r2} & \cdots & p_{rr} \end{bmatrix} = \left[\sum_{i=1}^r p_i p_{i1}, \sum_{i=1}^r p_i p_{i2}, \dots, \sum_{i=1}^r p_i p_{ir} \right]$$

STEADY STATE OR EQUILIBRIUM

The probability distribution of the outcomes for the n th observation of a first-order Markov process is the product of the initial probability vector and the n th power of the transition matrix. This can be shown by extending the arguments of the previous section. If the vector of initial probabilities is denoted by p_0 and the vector of probabilities at step n is denoted by p_n , then $p_1 = p_0 P$, $p_2 = p_1 P = p_0 P^2$, $p_3 = p_2 P = p_0 P^3$, ..., $p_n = p_0 P^n$. Under some rather nonrestrictive mathematical assumptions, it can be shown that a first-order Markov process approaches a steady state or equilibrium as the number of observations increases, that is, as n approaches infinity.

By definition, when a first-order Markov process is in equilibrium, the probability of each possible state or outcome is constant from observation to observation. The first-order Markov processes used to describe situations in business and economics almost invariably fulfill the mathematical assumptions for equilibrium, and the probability distribution of the states in equilibrium is frequently the most interesting property of these processes.

It can be shown that if a first-order Markov chain approaches a steady or equilibrium state as the number of observations or transitions approaches infinity, the stationary (equilibrium) probability distribution of its states is unique and depends only on the transition matrix and not on the initial probability distribution of the states. Any particular Markov chain is, of course, in one particular state for any given observation. The physical and computational significance of stationarity thus becomes apparent only if a number of Markov processes are in the same equilibrium state simultaneously. In this case if there are N such processes and p_i is the equilibrium probability of state i , $p_i N$ of the processes are expected to be in state i for any given observation. Thus, if N is large, a state of macroscopic equilibrium is maintained by a large number of transitions in opposite directions. Most statistical equilibria in physics are of this kind.

Computationally, the stationary probability distribution $[p_1, p_2, \dots, p_r]$ of the states of a Markov chain is obtained by solving the equations given in matrix form by

$$[p_1, p_2, \dots, p_r] \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1r} \\ p_{21} & p_{22} & \cdots & p_{2r} \\ \vdots & \vdots & & \vdots \\ p_{r1} & p_{r2} & \cdots & p_{rr} \end{bmatrix} = [p_1, p_2, \dots, p_r]$$

where p_i is the probability of being in state i , $i = 1, \dots, r$ and $\sum_{i=1}^r p_i = 1$. This involves solution of the r independent linear equations

$$\sum_{i=1}^r p_i p_{ij} = p_j \quad \text{for } j = 1, 2, \dots, r \quad (r-1 \text{ of which are independent})$$

$$\sum_{i=1}^r p_i = 1$$

Note that these equations actually define equilibrium: If $[p_1, p_2, \dots, p_r]$ is the probability distribution of the states for a given observation, the product

$$[p_1, p_2, \dots, p_r] \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1r} \\ p_{21} & p_{22} & \cdots & p_{2r} \\ \vdots & \vdots & & \vdots \\ p_{r1} & p_{r2} & \cdots & p_{rr} \end{bmatrix}$$

gives the probability distribution of the states for the next observation. If that distribution is also $[p_1, p_2, \dots, p_r]$, the same as for the preceding observation, the system is in equilibrium.

EXAMPLE

Suppose that, on September 1, of the subscribers in a fixed area, the Herald has $\frac{1}{2}$, the Tribune has $\frac{1}{4}$, and the Gazette has $\frac{1}{4}$. During the month of September, the Herald retains $\frac{7}{8}$ of its subscribers and loses $\frac{1}{8}$ of them to the Tribune; the Tribune retains $\frac{1}{12}$ of its subscribers and loses $\frac{3}{4}$ of them to the Herald and $\frac{1}{6}$ of them to the Gazette; the Gazette retains $\frac{1}{3}$ of its subscribers and loses $\frac{1}{2}$ of them to the Herald and $\frac{1}{6}$ of them to the Tribune. Assume there are no new subscribers and that none quit subscribing.

- What proportion of the subscribers does each paper have on October 1?
- If the same pattern of gains and losses continues for October, what proportion of the subscribers does each paper have on November 1?
- If the same pattern of gains and losses continues each month, what proportion of the subscribers will each paper have in the long run (that is, in equilibrium)?

The transition matrix is

	Herald	Tribune	Gazette
Herald	$\begin{bmatrix} \frac{7}{8} & \frac{1}{8} & 0 \end{bmatrix}$		
Tribune	$\begin{bmatrix} \frac{3}{4} & \frac{1}{12} & \frac{1}{6} \end{bmatrix}$		
Gazette	$\begin{bmatrix} \frac{1}{2} & \frac{1}{6} & \frac{1}{3} \end{bmatrix}$		

(a)

$$\begin{bmatrix} \frac{1}{2}, \frac{1}{4}, \frac{1}{4} \end{bmatrix} \begin{bmatrix} \frac{7}{8} & \frac{1}{8} & 0 \\ \frac{3}{4} & \frac{1}{12} & \frac{1}{6} \\ \frac{1}{2} & \frac{1}{6} & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{3}{4}, \frac{1}{8}, \frac{1}{8} \end{bmatrix}$$

Thus on October 1, of the subscribers, the Herald has $\frac{3}{4}$, the Tribune has $\frac{1}{8}$, and the Gazette has $\frac{1}{8}$.

(b)

$$\begin{bmatrix} \frac{3}{4}, \frac{1}{8}, \frac{1}{8} \end{bmatrix} \begin{bmatrix} \frac{7}{8} & \frac{1}{8} & 0 \\ \frac{3}{4} & \frac{1}{12} & \frac{1}{6} \\ \frac{1}{2} & \frac{1}{6} & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{13}{16}, \frac{1}{8}, \frac{1}{16} \end{bmatrix}$$

Thus on November 1, of the subscribers, the Herald has $\frac{13}{16}$, the Tribune has $\frac{1}{8}$, and the Gazette has $\frac{1}{16}$.

(c)

$$\begin{bmatrix} p_1, p_2, p_3 \end{bmatrix} \begin{bmatrix} \frac{7}{8} & \frac{1}{8} & 0 \\ \frac{3}{4} & \frac{1}{12} & \frac{1}{6} \\ \frac{1}{2} & \frac{1}{6} & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} p_1, p_2, p_3 \end{bmatrix}$$

$$\frac{7}{8}p_1 + \frac{3}{4}p_2 + \frac{1}{2}p_3 = p_1$$

$$\frac{1}{8}p_1 + \frac{1}{12}p_2 + \frac{1}{6}p_3 = p_2$$

$$\frac{1}{6}p_2 + \frac{1}{3}p_3 = p_3$$

$$p_1 + p_2 + p_3 = 1$$

Multiplying the first equation by 8, the second equation by 24, and the third equation by 6,

$$-p_1 + 6p_2 + 4p_3 = 0$$

$$3p_1 - 22p_2 + 4p_3 = 0$$

$$p_2 - 4p_3 = 0$$

$$p_1 + p_2 + p_3 = 1$$

Solving the first, third, and fourth equations simultaneously (any three of the four equations are independent),

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & -4 & 0 \\ -1 & 6 & 4 & 0 \end{array} \right]$$

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & -4 & 0 \\ 0 & 7 & 5 & 1 \end{array} \right]$$

$$\left[\begin{array}{ccc|c} 1 & 0 & 5 & 1 \\ 0 & 1 & -4 & 0 \\ 0 & 0 & 33 & 1 \end{array} \right]$$

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & \frac{28}{33} \\ 0 & 1 & 0 & \frac{4}{33} \\ 0 & 0 & 1 & \frac{1}{33} \end{array} \right]$$

Thus in the long run, of the subscribers, the Herald will have $\frac{28}{33}$, the Tribune will have $\frac{4}{33}$, and the Gazette will have $\frac{1}{33}$.

EXAMPLE

In the community of Gardenville, each year 5 percent of the residents in the city proper move to the suburbs and 2 percent of the people in the suburbs move to the city. Assuming that the total number of people in the community remains constant, determine the long-run proportions of city and suburban residents.

The transition matrix is

	City	Suburbs
City	0.95	0.05
Suburbs	0.02	0.98

and the equilibrium probabilities are determined by

$$[p_1, p_2] \begin{bmatrix} 0.95 & 0.05 \\ 0.02 & 0.98 \end{bmatrix} = [p_1, p_2]$$

$$0.95p_1 + 0.02p_2 = p_1$$

$$0.05p_1 + 0.98p_2 = p_2$$

$$p_1 + p_2 = 1$$

Multiplying the first two equations by 100,

$$-5p_1 + 2p_2 = 0$$

$$5p_1 - 2p_2 = 0$$

$$p_1 + p_2 = 1$$

Solving the first and third equations simultaneously (any two of the three equations are independent),

$$\left[\begin{array}{cc|c} 1 & 1 & 1 \\ -5 & 2 & 0 \end{array} \right]$$

$$\left[\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 7 & 5 \end{array} \right]$$

$$\left[\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 1 & \frac{5}{7} \end{array} \right]$$

$$\left[\begin{array}{cc|c} 1 & 0 & \frac{2}{7} \\ 0 & 1 & \frac{5}{7} \end{array} \right]$$

Thus, eventually, of the people in the community, $\frac{2}{7}$ will be city residents and $\frac{5}{7}$ will be residents of the suburbs. Note that in this equilibrium state each year $0.05 \times \frac{2}{7} = \frac{1}{70}$ of the people move from the city to the suburbs and $0.02 \times \frac{5}{7} = \frac{1}{70}$ of the people move from the suburbs to the city—the numbers of city and suburban residents thus are unchanged or stable.

EXAMPLE

A broker is studying the price movements of various stocks on the market and is particularly interested in a company called Astronaut Instruments. He has observed that if this stock goes up on a given day, then the next day it has a 50 : 50 chance of going up again, a $\frac{1}{3}$ chance of staying the same price, and a $\frac{1}{6}$ chance of going down. If the stock stays the same on a given day, then it is equally likely to go up, stay the same, or go down the next day. If the stock goes down on a given day, then the next day it has a 50 : 50 chance of going down again, a $\frac{1}{3}$ chance of staying the same price, and a $\frac{1}{6}$ chance of going up. What proportion of the time (in the long run) does the stock go up, stay the same, and go down?

The transition matrix is

	Up	Same	Down
Up	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{6}$
Same	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
Down	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{2}$

and the equilibrium probabilities are determined by

$$[p_1, p_2, p_3] \begin{bmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{2} \end{bmatrix} = [p_1, p_2, p_3]$$

$$\frac{1}{2}p_1 + \frac{1}{3}p_2 + \frac{1}{6}p_3 = p_1$$

$$\frac{1}{3}p_1 + \frac{1}{3}p_2 + \frac{1}{3}p_3 = p_2$$

$$\frac{1}{6}p_1 + \frac{1}{3}p_2 + \frac{1}{2}p_3 = p_3$$

$$p_1 + p_2 + p_3 = 1$$

Multiplying the first equation by 6, the second equation by 2, and the third equation by 6,

$$-3p_1 + 2p_2 + p_3 = 0$$

$$p_1 - 2p_2 + p_3 = 0$$

$$p_1 + 2p_2 - 3p_3 = 0$$

$$p_1 + p_2 + p_3 = 1$$

Solving the second, third, and fourth equations simultaneously (any three of the four equations are independent),

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & -3 & 0 \\ 1 & -2 & 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & -4 & -1 \\ 0 & -3 & 0 & -1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 5 & 2 \\ 0 & 1 & -4 & -1 \\ 0 & 0 & -12 & -4 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 & \frac{1}{3} \\ 0 & 1 & 0 & \frac{1}{3} \\ 0 & 0 & 1 & \frac{1}{3} \end{bmatrix}$$

Thus, in the long run, the stock goes up $\frac{1}{3}$ of the time, stays the same $\frac{1}{3}$ of the time, and goes down $\frac{1}{3}$ of the time.

PROBLEMS

1. A country has a three-party political system and the results of elections follow a definite pattern: If a party wins an election, its chance of winning the next election are 50 : 50 and if it loses the next election, each of the other two parties has a 50 : 50 chance of winning. What proportion of the elections does each party win over a long period of time?
2. A trucking company offers its drivers three approved routes between two cities: over the Mystic Bridge, on the Interstate Parkway, and on Route 1. If a trucker goes on the Mystic Bridge, the chance of his getting into a traffic jam is $\frac{1}{3}$; if he does get in a traffic jam, the next day he takes the Interstate Parkway with probability $\frac{2}{3}$ and Route 1 with probability $\frac{1}{3}$; if he doesn't get in a traffic jam, the next day he takes the Mystic Bridge again with probability $\frac{1}{2}$, the Interstate Parkway with probability $\frac{1}{6}$, and Route 1 with probability $\frac{1}{3}$. If he takes the Interstate Parkway the chance of his getting into a traffic jam is $\frac{1}{2}$; if he does get in a traffic jam, the next day he takes the Mystic Bridge; if he doesn't get in a traffic jam, the next day he takes the Mystic Bridge, the Interstate Parkway, and Route 1 with equal probability. If he takes Route 1, the trucker is invariably late, so he never takes Route 1 two days in a row and the next day he takes the Mystic Bridge with probability $\frac{1}{3}$ and the Interstate Parkway with probability $\frac{2}{3}$. What proportion of the time does the trucker go on the Mystic Bridge, the Interstate Parkway, and Route 1, respectively?
3. Every year the Smith family goes on a vacation—a camping trip (preferred by the children), a visit to the city (preferred by Mrs. Smith), or a winter vacation (preferred by Mr. Smith). They never take the same kind of vacation 2 years in a row. Each year they flip a coin to decide which of the two types of vacation they did not take the previous year they will take that year. What proportion of the time do the Smiths go on a camping trip, visit the city, and take a winter vacation, respectively?
4. A mechanic has a very unreliable automobile. Every morning he goes out to the garage hoping to start it; some days it starts by itself, some days it starts if he gets a push from a neighbor, other days no amount of pushing is effective and the service station must be called. If it starts one day, the chances are $\frac{1}{2}$ it will start, $\frac{1}{3}$ it must be pushed, and $\frac{1}{6}$ the service station must be called the next day. If it is pushed one day, the chances are $\frac{1}{3}$ it will start and $\frac{2}{3}$ the service station must be called the next day (it is never pushed 2 days in a row). If the service station is called one day, the chances are $\frac{2}{3}$ it will start and $\frac{1}{3}$ it must be pushed the next day (the service station is never needed 2 days in a row). In the long run, what proportion of the days will the auto start, be pushed, and need the service station, respectively?

5. The students of Professor Geology never know what's going to happen in class: the professor may give them a surprise quiz, take them on a field trip, discuss the daily assignment, or deliver a lecture. If he gives a quiz one day, the next day there is always a field trip. If there is a field trip one day, there is never a quiz or a field trip the next day, but discussing the assignment and a lecture are equally likely. If he discusses the assignment one day, there is $\frac{1}{4}$ chance of a quiz, $\frac{1}{4}$ chance of a field trip, $\frac{1}{8}$ chance of discussing the assignment, and $\frac{1}{8}$ chance of a lecture the next day. If he lectures one day, there is $\frac{1}{4}$ chance of a quiz, $\frac{1}{8}$ chance of a field trip, $\frac{1}{8}$ chance of discussing the assignment, and $\frac{1}{2}$ chance of a lecture the next day. One of the students has calculated that they can expect quizzes $\frac{43}{287}$ of the days, field trips $\frac{72}{287}$ of the days, discussion of the assignment $\frac{60}{287}$ of the days, and lectures $\frac{112}{287}$ of the days. Is he correct?
6. Mr. *S* (who prefers steak), Mr. *C* (who prefers chicken), and Mr. *H* (who prefers ham) are invited frequently to the home of a friend who serves only these three entrees. They have decided to bet before each dinner engagement on which entree will be served. Suppose that if the hostess serves steak on one occasion, she never serves it next time; she tosses a fair coin and serves chicken if it comes up heads and ham if it comes up tails. If she serves chicken on one occasion, then the next time she tosses two dice and serves chicken if the dice come up the same; if the dice come up different on the first toss, she tosses them again and serves steak if they come up the same and ham otherwise. If she serves ham on one occasion, then the next time she draws a card randomly from a deck and serves chicken if it is a diamond and steak otherwise. If each man always bets on his preference, what proportion of the time, in the long run, will each win?
7. Every summer the Stormy Lakes Yachting Association must decide whether to hold its annual regatta in June, July, or August. If the regatta is in June, the probability of good weather is $\frac{2}{3}$; if there is good weather, the next year the regatta will be held in June with probability $\frac{2}{3}$, in July with probability $\frac{1}{3}$, and in August with probability $\frac{1}{3}$; if there is bad weather, the next year the regatta will be held in July and August with equal probabilities. If the regatta is in July, good and bad weather are equally probable; if there is good weather, the next year the regatta will be held in July; if there is bad weather, the regatta will be held in August with probability $\frac{2}{3}$ and in June with probability $\frac{1}{3}$. If the regatta is in August, the probability of good weather is $\frac{2}{3}$; if there is good weather, the next year the regatta will be held in July and August with equal probabilities; if there is bad weather, the regatta will be held in June with probability $\frac{1}{3}$ and in July with probability $\frac{2}{3}$. What proportion of the time is the regatta held in June, July, and August, respectively?
8. A scientific book club has three mailing lists: a selected list consisting of members of 3 years or more of active membership, a membership list consisting of current members, and an augmented list consisting of current members and others who have shown interest in the club's activities in the past. For each new publication, it must be decided which list to use for mailing the sales literature. If after the preceding mailing there were too few copies ordered, the secretary uses the member list with probability $\frac{1}{4}$ and the augmented list with probability $\frac{3}{4}$; if after the preceding mailing there was a satisfactory response, the secretary uses the member list and the augmented list with equal probability; if after the preceding mailing, the publication was oversold, the secretary uses the selected list with probability $\frac{1}{3}$ and the member list with probability $\frac{2}{3}$. If the selected list is used, there are too few sold, or a satisfactory number sold with equal probability; if the member list is used, there are too few sold with probability $\frac{1}{4}$ and an adequate sale with probability $\frac{3}{4}$; if the augmented list is used, there is an adequate sale with probability $\frac{1}{3}$ and an oversale with probability $\frac{2}{3}$. What proportion of the time are the sales too small, adequate, and an oversale, respectively?
9. A housewife always buys one of three brands of detergent: *A*, *B*, or *C*. Which brand she buys depends partly on which, if any, of the three companies are having a promotional campaign (free combs, plastic roses, etc.). The companies time these campaigns at random, paying no attention to whether the competition is running one at the same time. Company *A* runs a campaign $\frac{1}{2}$ of the time, company *B* runs a campaign $\frac{1}{3}$ of the time, and company *C* runs a campaign $\frac{1}{3}$ of the time. If the housewife buys brand *A* one time, the next time she buys brand *A* if brand *A* is running a campaign or if none of the brands is running a campaign, she buys brand *B* if brand *B* is running a campaign but brand *A* is not, and she buys brand *C* if brand *C* is running the only campaign at the time. If she buys brand *B* one

time, the next time she buys brand A if it is running the only campaign at the time and brand B otherwise. If she buys brand C one time, the next time she buys brand A if it is running the only campaign at the time, brand C if it is running a campaign and B is not, and brand B otherwise. What proportion of the time does the housewife buy brands A , B , and C , respectively?

HELPFUL HINT: From $P(A) = \frac{1}{2}$, $P(B) = \frac{1}{3}$, $P(C) = \frac{1}{3}$ given above, the following probabilities can be obtained with a little arithmetic:

$$P(A \text{ only}) = \frac{2}{9} \quad P(A \text{ and } C, \text{ not } B) = \frac{1}{9}$$

$$P(B \text{ only}) = \frac{1}{9} \quad P(B \text{ and } C, \text{ not } A) = \frac{1}{18}$$

$$P(C \text{ only}) = \frac{1}{9} \quad P(A, B \text{ and } C) = \frac{1}{18}$$

$$P(A \text{ and } B, \text{ not } C) = \frac{1}{9} \quad P(\text{no campaign}) = \frac{2}{9}$$

10. Computeronics Corporation orders 2-foot lengths of wire for use in its experimental models from a supplier who claims that the wire has a specified strength desired by the Computeronics engineers. Each shipment is classified by the engineers who use it as satisfactory in strength (no report is made to the supplier), below standard (the supplier is notified by letter that the shipment was not as claimed), or unacceptable (the shipment is returned and a replacement which has been subjected to a 100% test is sent by the supplier in accordance with the guarantee). The classification depends upon the extent to which a shipment fulfills or fails to fulfill the supplier's claim regarding strength. The engineers have observed that if a shipment is classified as satisfactory, the next shipment is satisfactory $\frac{4}{5}$ of the time, below standard $\frac{3}{20}$ of the time, and unacceptable $\frac{1}{20}$ of the time; if a shipment is classified as below standard, the next shipment is satisfactory $\frac{19}{20}$ of the time, below standard $\frac{1}{20}$ of the time, and never unacceptable. What proportion of the shipments is satisfactory, below standard, and unacceptable in the long run, respectively?

Answers to Odd-Numbered Problems

1. $\frac{1}{3}, \frac{1}{3}, \frac{1}{3}$

3. $\frac{1}{3}, \frac{1}{3}, \frac{1}{3}$

5. yes

7. $\frac{24}{91}, \frac{42}{91}, \frac{25}{91}$

9. $\frac{4}{9}, \frac{31}{63}, \frac{4}{63}$

Algebra Review Problems

EXPONENTS

I. PROBLEMS

Write each of the following in radical form.

1. $k^{3/2}$
2. $(ab)^{3/4}$
3. $(m - n^2)^{-1/2}$
4. $6y^{1/5}$
5. $-4xy^{1/2}$
6. $(xy)^{2/3}$
7. $(x - y)^{5/6}$
8. $(a + 5y)^{1/2}$
9. $k^{1/2} - m^{1/2}$
10. $(4x - 6y)^{3/7}$
11. $y^{-2/3}$
12. $(a^2 - b^2)^{-1/2}$

Answers

1. $\sqrt{k^3}$
2. $\sqrt[4]{a^3b^3}$
3. $\frac{1}{\sqrt{m - n^2}}$
4. $6\sqrt[5]{y}$
5. $-4x\sqrt{y}$
6. $\sqrt[3]{x^2y^2}$
7. $\sqrt[6]{(x - y)^5}$
8. $\sqrt{a + 5y}$
9. $\sqrt{k} - \sqrt{m}$
10. $\sqrt[7]{(4x - 6y)^3}$
11. $\frac{1}{\sqrt[3]{y^2}}$
12. $\frac{1}{\sqrt{a^2 - b^2}}$

II. PROBLEMS

Write each of the following with positive fractional exponents.

1. $\sqrt{4^3}$
2. $\frac{1}{\sqrt{m - 1}}$
3. $\sqrt[3]{5x^2}$
4. $\sqrt[3]{ab}$
5. $\sqrt[5]{x^4}$
6. $\frac{x}{\sqrt[5]{y}}$
7. $x\sqrt[3]{a^2b^3}$
8. $\sqrt{x} - 5\sqrt{y}$
9. $\frac{2}{\sqrt{x + y}}$
10. $\frac{1}{\sqrt[3]{x}}$
11. $\sqrt[3]{x - y}$
12. $\sqrt[5]{7ab^2}$
13. $8\sqrt{m}$
14. $\frac{-6x}{\sqrt[3]{a^2b}}$
15. $-4\sqrt[5]{x^3y}$

Answers

1. $4^{3/2}$ 2. $\frac{1}{(m-1)^{1/2}}$ 3. $(5x^2)^{1/3} = 5^{1/3}x^{2/3}$ 4. $(ab)^{1/3}$ 5. $x^{4/5}$ 6. $\frac{x}{y^{1/5}}$
7. $xa^{2/3}b$ 8. $x^{1/2} - 5y^{1/2}$ 9. $\frac{2}{(x+y)^{1/2}}$ 10. $\frac{1}{x^{1/3}}$ 11. $(x-y)^{1/3}$
12. $(7ab^2)^{1/5} = 7^{1/5}a^{1/5}b^{2/5}$ 13. $8m^{1/2}$ 14. $\frac{-6x}{a^{2/3}b^{1/3}}$ 15. $-4(x^3y)^{1/5} = -4x^{3/5}y^{1/5}$

III. PROBLEMS

Determine each of the following indicated roots.

1. $\sqrt[5]{-32}$ 2. $\sqrt{2}$ 3. $\sqrt[4]{x^8y^4}$ 4. $-\sqrt[3]{27x^3}$ 5. $-\sqrt{16}$ 6. $-\sqrt{a^8y^4}$
7. $\sqrt[3]{-27}$ 8. $\sqrt[3]{\frac{27}{125}x^3y^9}$ 9. $-\sqrt{a^4b^6}$ 10. $\sqrt[3]{27a^6}$

Answers

1. -2 2. 3 3. x^2y 4. $-3x$ 5. -4 6. $-a^4y^2$ 7. -3
8. $\frac{3}{5}xy^3$ 9. $-a^2b^3$ 10. $3a^2$

IV. PROBLEMS

Change each of the following to simplest form.

1. $\sqrt{300}$ 2. $\sqrt{5xy}\sqrt{5x}$ 3. $\sqrt{2x^5y^2}$ 4. $-\sqrt{300}$ 5. $\sqrt{k^4}$ 6. $-\sqrt{m^3}$
7. $\sqrt{9x^5}$ 8. $-\sqrt{9x^4}$ 9. $-\sqrt{24x^6}$ 10. $\sqrt{8\sqrt{2}}$ 11. $\sqrt{ab}\sqrt{a^5b}$
12. $\sqrt[4]{a^3}\sqrt[4]{a}$ 13. $\sqrt[3]{3}\sqrt[3]{9}$ 14. $\sqrt{400}$ 15. $\sqrt{300}$ 16. $\sqrt[3]{a^7y}$

Answers

1. $\sqrt{100}\sqrt{3} = 10\sqrt{3}$ 2. $\sqrt{25x^2y} = \sqrt{25x^2}\sqrt{y} = 5x\sqrt{y}$ 3. $\sqrt{x^4y^2}\sqrt{2x} = x^2y\sqrt{2x}$
4. $-\sqrt{100}\sqrt{3} = -10\sqrt{3}$ 5. k^2 6. $-m\sqrt{m}$ 7. $3x\sqrt{x}$ 8. $-3x^2$
9. $-\sqrt{(4)(6)}x^6 = -2x\sqrt{6}$ 10. $\sqrt{(8)(2)} = 4$ 11. $\sqrt{a^6b^2} = a^3b$ 12. $\sqrt[4]{a^4} = a$
13. $\sqrt[3]{27} = 3$ 14. $\sqrt{4(10)^2} = 20$ 15. $\sqrt{3(10)^2} = 10\sqrt{3}$ 16. $a^2\sqrt[3]{ay}$

V. PROBLEMS

Rationalize the denominator in each of the following.

1. $\sqrt{\frac{1}{6}}$ 2. $\frac{\sqrt{6x}\sqrt{3x}}{\sqrt{9}}$ 3. $\frac{\sqrt[5]{3}}{a^3}$ 4. $\sqrt{\frac{1}{8}}$ 5. $\frac{-1}{\sqrt{3}}$ 6. $-\sqrt{\frac{e}{m}}$ 7. $\frac{-ab}{\sqrt{b}}$
8. $\frac{1}{\sqrt[3]{9}}$

Answers

1. $\frac{\sqrt{1}}{\sqrt{6}} = \frac{\sqrt{1}\sqrt{6}}{\sqrt{6}\sqrt{6}} = \frac{\sqrt{6}}{6}$ 2. $\frac{\sqrt{18x^2}}{\sqrt{9}} = \sqrt{2x^2} = x\sqrt{2}$ 3. $\frac{\sqrt{3}}{\sqrt[5]{a^3}} = \frac{\sqrt[5]{3}\sqrt[5]{a^2}}{\sqrt[5]{a^3}\sqrt[5]{a^2}} = \frac{\sqrt[5]{3a^2}}{a}$
4. $\frac{\sqrt{1}}{\sqrt{8}} \cdot \frac{\sqrt{2}}{\sqrt{2}} = \frac{\sqrt{2}}{4}$ 5. $-\frac{1}{\sqrt{3}} \cdot \frac{\sqrt{3}}{\sqrt{3}} = -\frac{\sqrt{3}}{3}$ 6. $\frac{-\sqrt{e}\sqrt{m}}{\sqrt{m}\sqrt{m}} = \frac{-\sqrt{em}}{m}$
7. $\frac{-ab\sqrt{b}}{\sqrt{b}\sqrt{b}} = -a\sqrt{b}$ 8. $\frac{1\sqrt[3]{3}}{\sqrt[3]{9}\sqrt[3]{3}} = \frac{\sqrt[3]{3}}{2}$

VI. PROBLEMS

Simplify each of the following.

1. $(2a^2b)(2ab^2)$
2. $(-3x^3y^4)(7y^3xz)$
3. $(x^{1/2})(yx^{1/2})$
4. $\frac{x^4}{x^7}$
5. $\frac{8a^4b^2}{4a^2b}$
6. $\frac{12x^3y^3z^3}{-6xyz}$
7. $\frac{48x^{3/2}y^{3/5}}{16x^{1/2}y^{1/5}}$
8. $(3x^2)^2$
9. $(3ab^{1/2})^2$
10. $(x^{2/3}y^{4/3})^3$
11. $(x^{2/3}y^{4/3})^{3/2}$
12. $\frac{(x^{1/2} \cdot y^4)^2}{(x^{2/3} \cdot y)^3}$
13. $\frac{(x^n)^{3/2}}{x^{n/2}}$
14. $(y^k \cdot y^{k/2})^4$
15. $(x^2)^{n/2} \cdot (y^{2n})^{2/n}$
16. $\left(\frac{x^n}{y}\right)^{1/2} \cdot \left(\frac{y}{x^{2n}}\right)^{3/2}$
17. $\left(\frac{a^mb^{2m}}{b^{4m}}\right)^{1/m}$
18. $x^4(8a^6x^3)^{-1/3}$
19. $\frac{4a^3b}{(a^4b^6)^{1/2}}$
20. $\frac{(x^{3/m}y^{2/m})^m}{x^my^{1-m}}$
21. $\frac{6a^{-3}x^2}{3(y^2c^{-4})^{1/2}}$
22. $(a^2b^{-3})^{-1}$
23. $x^0 \cdot y^n$
24. $\frac{4^{-1}a \cdot b^{-3}}{(2ab)^{-4}}$
25. $\frac{x^{-2}y^{-1}}{3y}$
26. $a^{-1} + b^{-2}$
27. $(n^{-1} + n^{-2})^{-1}$
28. $\left(\frac{n^{1-a}}{n^{2-a}}\right)^{-2}$
29. $(x)(x^{2/3})$
30. $n(n^{-3/5} - n^{-2/5})$
31. $(a+b)^{1/2}[(a+b)^{1/2} - (a+b)]$
32. $a^{-1/2}(a^{3/2} - a + 6)$

Answers

1. $4a^3b^3$
2. $-21x^4y^7z$
3. xy
4. $x^{-3} = \frac{1}{x^3}$
5. $2a^2b$
6. $-2x^2y^2z^2$
7. $3xy^{2/5}$
8. $9x^4$
9. $9a^2b$
10. x^2y^4
11. xy^2
12. $\frac{y^5}{x}$
13. x^h
14. y^{6k}
15. x^ny^4
16. $\frac{y}{x^{5n/2}}$
17. $\frac{a}{b^2}$
18. $\frac{x^3}{2a^2}$
19. $\frac{4a}{b^2}$
20. $x^{3-m}y^{1+m}$
21. $\frac{2c^2x^2}{a^3y}$
22. $\frac{b^3}{a^2}$
23. y^n
24. $4a^5b$
25. $\frac{1}{3x^2y^2}$
26. $\frac{b^2+a}{ab^2}$
27. $\frac{n^2}{n+1}$
28. n^2
29. $x^{5/3}$
30. $n^{2/5} - n^{3/5}$
31. $(a+b) - (a+b)^{3/2}$
32. $a - a^{1/2} + 6a^{-1/2}$

LOGARITHMS

I. PROBLEMS

Write each of the following as the sum or difference of simpler logarithmic quantities.

1. $\log_a\left(\frac{mn}{x}\right)^{1/3}$
2. $\log_x(abc)$
3. $\log_n\left(\frac{a}{b}\right)$
4. $\log_x\sqrt[5]{y}$
5. $\log_a n^5$
6. $\log_a x^5y^2$
7. $\log_a\sqrt{\frac{m}{n}}$
8. $\log_4\sqrt{a}\sqrt[5]{b^2}$
9. $\log_3\sqrt{a(a-b)}$
10. $\log_{10}\sqrt[5]{\frac{ab^2}{c}}$

Answers

1. $\frac{1}{3}(\log_a m + \log_a n - \log_a x)$ 2. $\log_x a + \log_x b + \log_x c$ 3. $\log_n a - \log_n b$
 4. $\frac{1}{3} \log_x y$ 5. $5 \log_a n$ 6. $5 \log_a x + 2 \log_a y$ 7. $\frac{1}{2} \log_a m - \frac{1}{2} \log_a n$
 8. $\frac{1}{2} \log_4 a + \frac{2}{3} \log_4 b$ 9. $\frac{1}{2} \log_3 a + \frac{1}{2} \log_3 (a - b)$ 10. $\frac{1}{3} \log_{10} a + \frac{2}{3} \log_{10} b - \frac{1}{3} \log_{10} c$

EXPONENTIAL FUNCTIONS

I. PROBLEMS

1. Find the second component of each of the following ordered pairs (x, y) if $y = 2^x$:
 $(-3, \quad)$, $(0, \quad)$, $(3, \quad)$, $(6, \quad)$. 2. Find the second component of each of the following
 ordered pairs $(x, f(x))$ if $f(x) = (\frac{1}{2})^x$: $(-3, \quad)$, $(0, \quad)$, $(3, \quad)$, $(5, \quad)$. 3. State the equations
 $a = b^x$ if $b = 10, e, \frac{1}{10}, \frac{1}{e}$.

Answers

1. $y = 2^{-3} = \frac{1}{8}$; $y = 2^0 = 1$; $y = 2^3 = 8$; $y = 2^6 = 64$
 2. $f(x) = (\frac{1}{2})^{-3} = 8$; $f(x) = (\frac{1}{2})^0 = 1$; $f(x) = (\frac{1}{2})^3 = \frac{1}{8}$; $f(x) = (\frac{1}{2})^5 = \frac{1}{32}$
 3. $a = 10^x$; $a = e^x$; $a = (\frac{1}{10})^x = 10^{-x}$; $a = (\frac{1}{e})^x = e^{-x}$

EXPONENTIAL AND LOGARITHMIC FUNCTIONS

I. PROBLEMS

Express each of the following exponential statements in logarithmic notation.

1. $5^2 = 25$ 2. $3^{-2} = \frac{1}{9}$ 3. $16^{1/4} = 2$ 4. $3^2 = 9$ 5. $2^3 = 8$ 6. $(\frac{1}{3})^2 = \frac{1}{9}$
 7. $10^2 = 100$ 8. $10^3 = 1000$ 9. $10^{-1} = 0.1$ 10. $10^0 = 1$

Answers

1. $\log_5 25 = 2$ 2. $\log_3 \frac{1}{9} = -2$ 3. $\log_{16} 2 = \frac{1}{4}$ 4. $\log_3 9 = 2$
 5. $\log_2 8 = 3$ 6. $\log_{1/3} \frac{1}{9} = 2$ 7. $\log_{10} 100 = 2$ 8. $\log_{10} 1000 = 3$
 9. $\log_{10} 0.1 = -1$ 10. $\log_{10} 1 = 0$

II. PROBLEMS

Express each of the following logarithmic statements in exponential notation.

1. $\log_{1/5} 125 = -3$ 2. $\log_4 16 = 2$ 3. $\log_3 27 = 3$ 4. $\log_6 36 = 2$
 5. $\log_{10} 1000 = 3$ 6. $\log_{1/3} \frac{1}{9} = 2$ 7. $\log_{12} 144 = 2$ 8. $\log_e e = 1$

Answers

1. $(\frac{1}{5})^{-3} = 125$ 2. $4^2 = 16$ 3. $3^3 = 27$ 4. $6^2 = 36$ 5. $10^3 = 1000$
 6. $(\frac{1}{3})^2 = \frac{1}{9}$ 7. $12^2 = 144$ 8. $e^1 = e$

III. PROBLEMS

Find the value of x , y , or z for each of the following.

1. $\log_4 16 = x$ 2. $\log_6 216 = y$ 3. $\log_{10} 1 = x$ 4. $\log_4 4 = z$
 5. $\log_{10} 100 = y$ 6. $\log_{10} 0.1 = y$ 7. $\log_{10} 0.01 = x$ 8. $\log_4 \frac{1}{4} = z$
 9. $\log_3 27 = x$ 10. $\log_{12} 144 = y$ 11. $\log_4 x = 2$ 12. $\log_3 x = 2$
 13. $\log_2 4 = y$ 14. $\log_6 36 = z$ 15. $\log_x 10 = \frac{1}{2}$ 16. $\log_3 27 = y$

Answers

- | | | | | | |
|-------------|-------------|-------------|---------------|--------------|-------------|
| 1. $x = 2$ | 2. $y = 3$ | 3. $x = 0$ | 4. $z = 1$ | 5. $y = 2$ | 6. $y = -1$ |
| 7. $x = -2$ | 8. $z = -1$ | 9. $x = 3$ | 10. $y = 2$ | 11. $x = 16$ | |
| 12. $x = 9$ | 13. $y = 2$ | 14. $z = 2$ | 15. $x = 100$ | 16. $y = 3$ | |

Tables

Exponential Functions

x	e^x	e^{-x}	x	e^x	e^{-x}
0.00	1.000	1.000	3.00	20.086	0.050
0.10	1.105	0.905	3.10	22.198	0.045
0.20	1.221	0.819	3.20	24.533	0.041
0.30	1.350	0.741	3.30	27.113	0.037
0.40	1.492	0.670	3.40	29.964	0.033
0.50	1.649	0.607	3.50	33.115	0.030
0.60	1.822	0.549	3.60	36.598	0.027
0.70	2.014	0.497	3.70	40.447	0.025
0.80	2.226	0.449	3.80	44.701	0.022
0.90	2.460	0.407	3.90	49.402	0.020
1.00	2.718	0.368	4.00	54.598	0.018
1.10	3.004	0.333	4.10	60.340	0.017
1.20	3.320	0.301	4.20	66.686	0.015
1.30	3.669	0.273	4.30	73.700	0.014
1.40	4.055	0.247	4.40	81.451	0.012
1.50	4.482	0.223	4.50	90.017	0.011
1.60	4.953	0.202	4.60	99.484	0.010
1.70	5.474	0.183	4.70	109.95	0.009
1.80	6.050	0.165	4.80	121.51	0.008
1.90	6.686	0.150	4.90	134.29	0.007
2.00	7.389	0.135	5.00	148.41	0.007
2.10	8.166	0.122	5.10	164.02	0.006
2.20	9.025	0.111	5.20	181.27	0.006
2.30	9.974	0.100	5.30	200.34	0.005
2.40	11.023	0.091	5.40	221.41	0.005
2.50	12.182	0.082	5.50	244.69	0.004
2.60	13.464	0.074	5.60	270.43	0.004
2.70	14.880	0.067	5.70	298.87	0.003
2.80	16.445	0.061	5.80	330.30	0.003
2.90	18.174	0.055	5.90	365.04	0.003
3.00	20.086	0.050	6.00	403.43	0.002

Common Logarithms

<i>N</i>	0	1	2	3	4	5	6	7	8	9
10	0000	0043	0086	0128	0170	0212	0253	0294	0334	0374
11	0414	0453	0492	0531	0569	0607	0645	0682	0719	0755
12	0792	0828	0864	0899	0934	0969	1004	1038	1072	1106
13	1139	1173	1206	1239	1271	1303	1335	1367	1399	1430
14	1461	1492	1523	1553	1584	1614	1644	1673	1703	1732
15	1761	1790	1818	1847	1875	1903	1931	1959	1987	2014
16	2041	2068	2095	2122	2148	2175	2201	2227	2253	2279
17	2304	2330	2355	2380	2405	2430	2455	2480	2504	2529
18	2553	2577	2601	2625	2648	2672	2695	2718	2742	2765
19	2788	2810	2833	2856	2878	2900	2923	2945	2967	2989
20	3010	3032	3054	3075	3096	3118	3139	3160	3181	3201
21	3222	3243	3263	3284	3304	3324	3345	3365	3385	3404
22	3424	3444	3464	3483	3502	3522	3541	3560	3579	3598
23	3617	3636	3655	3674	3692	3711	3729	3747	3766	3784
24	3802	3820	3838	3856	3874	3892	3909	3927	3945	3962
25	3979	3997	4014	4031	4048	4065	4082	4099	4116	4133
26	4150	4166	4183	4200	4216	4232	4249	4265	4281	4298
27	4314	4330	4346	4362	4378	4393	4409	4425	4440	4456
28	4472	4487	4502	4518	4533	4548	4564	4579	4594	4609
29	4624	4639	4654	4669	4683	4698	4713	4728	4742	4757
30	4771	4786	4800	4814	4829	4843	4857	4871	4886	4900
31	4914	4928	4942	4955	4969	4983	4997	5011	5024	5038
32	5051	5065	5079	5092	5105	5119	5132	5145	5159	5172
33	5185	5198	5211	5224	5237	5250	5263	5276	5289	5302
34	5315	5328	5340	5353	5366	5378	5391	5403	5416	5428
35	5441	5453	5465	5478	5490	5502	5514	5527	5539	5551
36	5563	5575	5587	5599	5611	5623	5635	5647	5658	5670
37	5682	5694	5705	5717	5729	5740	5752	5763	5775	5786
38	5798	5809	5821	5832	5843	5855	5866	5877	5888	5899
39	5911	5922	5933	5944	5955	5966	5977	5988	5999	6010
40	6021	6031	6042	6053	6064	6075	6085	6096	6107	6117
41	6128	6138	6149	6160	6170	6180	6191	6201	6212	6222
42	6232	6243	6253	6263	6274	6284	6294	6304	6314	6325
43	6335	6345	6355	6365	6375	6385	6395	6405	6415	6425
44	6435	6444	6454	6464	6474	6484	6493	6503	6513	6522
45	6532	6542	6551	6561	6571	6580	6590	6599	6609	6618
46	6628	6637	6646	6656	6665	6675	6684	6693	6702	6712
47	6721	6730	6739	6749	6758	6767	6776	6785	6694	6803
48	6812	6821	6830	6839	6848	6857	6866	6875	6884	6893
49	6902	6911	6920	6928	6937	6946	6955	6964	6972	6981
50	6990	6998	7007	7016	7024	7033	7042	7050	7059	7067
51	7076	7084	7093	7101	7110	7118	7126	7135	7143	7152
52	7160	7168	7177	7185	7193	7202	7210	7218	7226	7235
53	7243	7251	7259	7267	7275	7284	7292	7300	7308	7316
54	7324	7332	7340	7348	7356	7364	7372	7380	7388	7396

Common Logarithms

<i>N</i>	0	1	2	3	4	5	6	7	8	9
55	7404	7412	7419	7427	7435	7443	7451	7459	7466	7474
56	7482	7490	7497	7505	7513	7520	7528	7536	7543	7551
57	7559	7566	7574	7582	7589	7597	7604	7612	7619	7627
58	7634	7642	7649	7657	7664	7672	7679	7686	7694	7701
59	7709	7716	7723	7731	7738	7745	7752	7760	7767	7774
60	7782	7789	7796	7803	7810	7818	7825	7832	7839	7846
61	7853	7860	7868	7875	7882	7889	7896	7903	7910	7917
62	7924	7931	7938	7945	7952	7959	7966	7973	7980	7987
63	7993	8000	8007	8014	8021	8028	8035	8041	8048	8055
64	8062	8069	8075	8082	8089	8096	8102	8109	8116	8122
65	8129	8136	8142	8149	8156	8162	8169	8176	8182	8189
66	8195	8202	8209	8215	8222	8228	8235	8241	8248	8254
67	8261	8267	8274	8280	8287	8293	8299	8306	8312	8319
68	8325	8331	8338	8344	8351	8357	8363	8370	8376	8382
69	8388	8395	8401	8407	8414	8420	8426	8432	8439	8445
70	8451	8457	8463	8470	8476	8482	8488	8494	8500	8506
71	8513	8519	8525	8531	8537	8543	8549	8555	8561	8567
72	8573	8579	8585	8591	8597	8603	8609	8615	8621	8627
73	8633	8639	8645	8651	8657	8663	8669	8675	8681	8686
74	8692	8698	8704	8710	8716	8722	8727	8733	8739	8745
75	8751	8756	8762	8768	8774	8779	8785	8791	8797	8802
76	8808	8814	8820	8825	8831	8837	8842	8848	8854	8859
77	8865	8871	8876	8882	8887	8893	8899	8904	8910	8915
78	8921	8927	8932	8938	8943	8949	8954	8960	8965	8971
79	8976	8982	8987	8993	8998	9004	9009	9015	9020	9025
80	9031	9036	9042	9047	9053	9058	9063	9069	9074	9079
81	9085	9090	9096	9101	9106	9112	9117	9122	9128	9133
82	9138	9143	9149	9154	9159	9165	9170	9175	9180	9186
83	9191	9196	9201	9206	9212	9217	9222	9227	9232	9238
84	9243	9248	9253	9258	9263	9269	9274	9279	9284	9289
85	9294	9299	9304	9309	9315	9320	9325	9330	9335	9340
86	9345	9350	9355	9360	9365	9370	9375	9380	9385	9390
87	9395	9400	9405	9410	9415	9420	9425	9430	9435	9440
88	9445	9450	9455	9460	9465	9469	9474	9479	9484	9489
89	9494	9499	9504	9509	9513	9518	9523	9528	9533	9538
90	9542	9547	9552	9557	9562	9566	9571	9576	9581	9586
91	9590	9595	9600	9605	9609	9614	9619	9624	9628	9633
92	9638	9643	9647	9652	9657	9661	9666	9671	9675	9680
93	9685	9689	9694	9699	9703	9708	9713	9717	9722	9727
94	9731	9736	9741	9745	9750	9754	9759	9763	9768	9773
95	9777	9782	9786	9791	9795	9800	9805	9809	9814	9818
96	9823	9827	9832	9836	9841	9845	9850	9854	9859	9863
97	9868	9872	9877	9881	9886	9890	9894	9899	9903	9908
98	9912	9917	9921	9926	9930	9934	9939	9943	9948	9952
99	9956	9961	9965	9969	9974	9978	9983	9987	9991	9996

Selected References

MATHEMATICS

- Andree, Richard V. *Introduction to Calculus with Analytic Geometry*. New York: McGraw-Hill, 1962.
- Birkhoff, G., and MacLane, S. *A Survey of Modern Algebra*. 3d ed. New York: Macmillan, 1965.
- Britton, J. R., Kriegh, R. B., and Rutland, Leon W. *Calculus and Analytic Geometry*. San Francisco, Calif.: Freeman, 1966.
- Buck, R. C. *Advanced Calculus*. 2d ed. New York: McGraw-Hill, 1965.
- Burkill, J. C. *The Theory of Ordinary Differential Equations*. New York: Wiley, Halstead Press, 1956.
- Coddington, E. A., and Levinson, N. *Theory of Ordinary Differential Equations*. New York: McGraw-Hill, 1955.
- Courant, R. *Differential and Integral Calculus*. Translated by E. J. McShane. Vols. I and II. New York: Interscience, 1955 printing.
- Cullen, Charles. *Matrices and Linear Transformations*. 2d ed. Reading, Mass.: Addison-Wesley, 1972.
- Garvin, W. W. *Introduction to Linear Programming*. New York: McGraw-Hill, 1960.
- Glicksman, A. M. *Linear Programming and The Theory of Games*. New York: Wiley, 1963.
- Goldberg, S. *Introduction to Difference Equations*. New York: Wiley, 1958.
- Golomb, M., and Shanks, M. *Elements of Ordinary Differential Equations*. 2d ed. New York: McGraw-Hill, 1965.
- Hadley, G. *Linear Algebra*. Reading, Mass.: Addison-Wesley, 1961.
- Hadley, G. *Linear Programming*. Reading, Mass.: Addison-Wesley, 1962.
- Ince, E. L. *Ordinary Differential Equations*. New York: Dover, 1953.
- Jolley, L. B. W. *Summation of Series*, 2d ed. New York: Dover, 1961.
- Kaplan, W. *Advanced Calculus*. 2d ed. Reading, Mass.: Addison-Wesley, 1973.
- Kemeny, J. G. Snell, J. L., and Thompson, G. L. *Finite Markov Chains*. Englewood Cliffs, N.J.: Prentice-Hall, 1960.
- Kuhn, H. W., and Tucker, A. W., eds. *Linear Inequalities and Related Systems*. Princeton, N.J.: Princeton University Press, 1956.
- Lang, S. *A First Course in Calculus*. Reading, Mass.: Addison-Wesley, 1964.
- Llewellyn, R. W. *Linear Programming*. New York: Holt, Rinehart & Winston, 1964.
- Luce, R. D., and Raiffa, H. *Games and Decisions*. New York: Wiley, 1957.
- McKinsey, J. C. C. *Introduction to the Theory of Games*. New York: McGraw-Hill, 1952.
- Mirsky, L. *An Introduction to Linear Algebra*. 2d ed. New York: Oxford University Press, 1961.

- Nelson, A. L., Folley, K. W., and Coral, M. *Differential Equations*. 3d ed. Boston: D. C. Heath, 1964.
- Nemhauser, G. L. *Introduction to Dynamic Programming*. New York: Wiley, 1966.
- Niven, Ivan. *Calculus, An Introductory Approach*. 2d ed. Princeton, N.J.: D. Van Nostrand, 1966.
- Orchard-Hays, W. *Advanced Linear Programming Computer Techniques*. New York: McGraw-Hill, 1968.
- Perlis, S. *Theory of Matrices*. Reading, Mass.: Addison-Wesley, 1952.
- Peterson, T. S. *Elements of Calculus*. 2d ed. New York: Harper & Row, 1960.
- Richardson, C. H. *An Introduction to the Calculus of Finite Differences*. New York: Van Nostrand, Reinhold, 1954.
- Sasaki, Kyohei. *Introduction to Finite Mathematics and Linear Programming*. Belmont, Calif.: Wadsworth, 1970.
- Sisam, C. H., and Atchison, W. F. *Analytic Geometry*. New York: Holt, Rinehart & Winston, 1955.
- Smail, L. L. *Calculus*. New York: Appleton-Century-Crofts, 1949.
- Spitzbart, A., and Bardell, R. H. *College Algebra and Plane Trigonometry*. 2d. ed. Reading, Mass.: Addison-Wesley, 1964.
- Steen, F. H., and Ballou, D. H. *Analytic Geometry*. 3d ed. Lexington, Mass.: Xerox College Publishing, 1963.
- Stoll, R. F. *Linear Algebra and Matrix Theory*. New York: Dover, 1969.
- Thomas, G. B., Jr. *Calculus and Analytic Geometry*. 4th ed. Reading, Mass.: Addison-Wesley, 1968.
- Thrall, R. M., and Tornheim, L. *Vector Spaces and Matrices*. New York: Dover, 1970.
- Whipkey, K. L., and Whipkey, M. N. *The Power of Calculus*. 2d ed. New York: Wiley, 1975.
- Widder, D. V. *Advanced Calculus*. 2d ed. Englewood Cliffs, N.J.: Prentice-Hall, 1961.
- Wooton, W., and Drooyan, I. *Intermediate Algebra*. 3d ed. Belmont, Calif.: Wadsworth, 1972.
- Wylie, C. R., Jr. *Calculus*. New York: McGraw-Hill, 1953.

APPLICATIONS OF MATHEMATICS IN BUSINESS AND ECONOMICS

- Arrow, K., and others. *Studies in Linear and Non-Linear Programming*. Stanford, Calif.: Stanford University Press, 1958.
- Arrow, K., and Karlin, S. *Studies in the Mathematical Theory of Inventory and Production*. Stanford, Calif.: Stanford University Press, 1958.
- Beatty, William E. *Mathematical Relationships in Business and Economics*. Morristown, N.J.: General Learning Corporation, 1970.
- Beatty, William E. and Gage, R. William. *Introductory Calculus for Business and Economics*. Morristown, N.J.: General Learning Corporation, 1973.
- Bierman, H., Bonini, C. P., Jr., and Hausman, W. H. *Quantitative Analysis for Business Decisions*. 4th ed. Homewood, Ill.: Irwin, 1973.
- Bowen, E. K. *Mathematics—With Applications in Management and Economics*. 3d ed. Homewood, Ill.: Irwin, 1972.
- Buffa, E. S. *Production Inventory System: Planning and Control*. Rev. ed. Homewood, Ill.: Irwin, 1972.
- Carr, C. R., and Howe, C. W. *Introduction to Quantitative Decision Procedures in Management and Economics*. New York: McGraw-Hill, 1964.
- Chiang, Alpha C. *Fundamental Methods of Mathematical Economics*. New York: McGraw-Hill, 1974.
- Churchman, C. W., Ackoff, R. L., and Arnoff, E. L. *Introduction to Operations Research*. New York: Wiley, 1957.
- Daus, P. H., and Whyburn, W. M. *Algebra With Applications to Business and Economics*. Reading, Mass.: Addison-Wesley, 1961.
- Daus, P. H., and Whyburn, W. H. *Introduction to Mathematical Analysis*. Reading, Mass.: Addison-Wesley, 1958.
- Dorfman, R., Samuelson, P. A., and Solow, R. *Linear Programming and Economic Analysis*. New York: McGraw-Hill, 1958.

- Fowler, F. P., Jr., and Sandberg, E. W. *Basic Mathematics for Administration*. New York: Wiley, 1962.
- Horowitz, I. *An Introduction to Quantitative Business Analysis*. 2d ed. New York: McGraw-Hill, 1972.
- Horst, P. *Matrix Algebra for Social Scientists*. New York: Holt, Rinehart & Winston, 1963.
- Howell, J. E., and Teichroew, D. *Mathematical Analysis for Business Decisions*. Homewood, Ill.: Irwin, 1963.
- Intriligator, M. D. *Mathematical Optimization and Economic Theory*. Englewood Cliffs, N.J.: Prentice-Hall, 1971.
- Karlin, S. *Mathematical Methods and Theory in Games, Programming and Economics*. Reading, Mass.: Addison-Wesley, 1959.
- Kemeny, J. G., Snell, J. L., and Thompson, G. L. *Introduction to Finite Mathematics*. 2d ed. Englewood Cliffs, N.J.: Prentice-Hall, 1966.
- Kooros, A. *Elements of Mathematical Economics*. Boston: Houghton Mifflin, 1965.
- Levin, R. I., and Kirkpatrick, C. A. *Quantitative Approaches to Management*. 3d ed. New York: McGraw-Hill, 1971.
- Naddor, E. *Inventory Systems*. New York: Wiley, 1966.
- Sasieni, M., Yaspan, A., and Friedman, L. *Operations Research—Methods and Problems*. New York: Wiley, 1959.
- Teichroew, D. *An Introduction to Management Science, Deterministic Models*. New York: Wiley, 1964.
- Tintner, G., and Millham, C. B. *Mathematics and Statistics for Economists*. 2d ed. New York: Holt, Rinehart & Winston, 1970.

ECONOMICS

- Ackley, G. *Macroeconomic Theory*. New York: Macmillan, 1961.
- Allen, C. L. *Elementary Mathematics of Price Theory*. Belmont, Calif.: Wadsworth, 1962.
- Baumol, W. J. *Economic Dynamics*. 3d ed. New York: Macmillan, 1970.
- Beach, E. F. *Economic Models*. New York: Wiley, 1957.
- Boulding, K. E. *Economic Analysis*, 4th ed. New York: Harper & Row, 1966.
- Cyert, R. M., and March, J. G. *A Behavioral Theory of the Firm*. Englewood Cliffs, N.J.: Prentice-Hall, 1963.
- Dernburg, T. F., and Dernburg, J. D. *Macroeconomic Analysis*. Reading, Mass.: Addison-Wesley, 1969.
- Dernburg, T. F., and McDougall, D. M. *Macroeconomics*. New York: McGraw-Hill, 1968.
- Gale, David. *The Theory of Linear Economic Models*. New York: McGraw-Hill, 1960.
- Henderson, J. M., and Quandt, R. E. *Microeconomic Theory, A Mathematical Approach*, 2d ed. New York: McGraw-Hill, 1971.
- Hibdon, James E. *Price and Welfare Theory*. New York: McGraw-Hill, 1969.
- Keynes, J. M. *The General Theory of Employment Interest and Money*. New York: Harcourt, Brace, Jovanovich, 1935.
- Neumann, J. von, and Morgenstern, O. *Theory of Games and Economic Behavior*. 3d ed. Princeton, N.J.: Princeton University Press, 1954.
- Samuelson, P. A. *Foundations of Economic Analysis*. Cambridge, Mass.: Harvard University Press, 1947.
- Samuelson, P. A. *Economics: Introductory Analysis*, 9th ed. New York: McGraw-Hill, 1973.
- Schultz, H. *The Theory and Measurement of Demand*. Chicago: University of Chicago Press, 1957.
- Shapiro, Edward. *Macroeconomics: Selected Readings*. New York: Harcourt, Brace and World, 1970.
- Stonier, A. W., and Hague, D. C. *A Textbook of Economic Theory*. 4th ed. New York: Wiley, Halstead Press, 1973.

Index

- Abscissa, 16. *See also*
Rectangular coordinates
- Adjoint of a matrix, 566, 577, 647–650. *See also*
Determinant of a matrix;
Inverse of a matrix;
Matrix
- Algebraic function, 50, 174–177, 201, 317–318. *See also*
Function
- Alternating series, 299–300. *See also* Infinite series
- Angle of inclination, 19–20
- Antiderivative, 388
- Approximate integration, 449–458. *See also* Integration, approximate methods of
- Approximation, 204–206. *See also* Differential
- Area. *See also* Integration; Probability beneath a curve, 404–408, 414–417
between two curves, 408–413
negative, 407–408
- Asymptote, 57–59
horizontal, 58–59
vertical, 58–59
- Average cost, 168, 245–246, 394–395. *See also* Cost
- Average revenue, 170, 266–267, 395–396. *See also*
Revenue
- Base e , 114–115. *See also*
Logarithm; Logarithmic function
- Bernoulli equations, 498
- Bilinear form, 597–598. *See also*
Vector differentiation
- Bivariate regression analysis, 363–364. *See also* Least squares estimates;
Regression analysis
- Bordered Hessian determinant, 607. *See also* Hessian determinant; Maximum of a function of n variables
- Boundary conditions. *See also*
Constant of integration;
Initial conditions;
Particular solution
for a difference equation, 512
for a differential equation, 472
for integration, 389
- Break-even analysis, 44–46
- Capital formation, 397–398. *See also* Integration; Net investment
- Cauchy's test ratio, 299–300. *See also* Convergence of an infinite series
- Circle, 72–74. *See also* Conic section
center of, 72
radius of, 72–73
standard form of, 72–73
- Cobweb model, 524
- Cofactor of a matrix, 566. *See also* Determinant of a matrix; Matrix
- Column vector, 542. *See also* Vector
- Competitive commodities, 341–344. *See also*
Complementary commodities; Demand surface
- Complementary commodities, 341–344. *See also*
Competitive commodities;
Demand surface
- Completing the square, 72, 134–137
- Composite function, 10, 196–197, 202, 321, 334–335. *See also* Function
- Compound interest, 118, 521. *See also* Exponential function; Interest; Simple interest
- Concave function, 611
- Concavity, 231–236, 240. *See also* Convexity; Point of inflection; Second derivative
concave downward, 231–232, 240
concave upward, 231–232, 240
- Conic section, 71–90. *See also*
Circle; Ellipse;
Hyperbola; Parabola;
Quadratic equation
circular, 72–74
degenerate, 71, 73, 75, 79, 83
directrix of, 71
eccentricity of, 71
elliptic, 72, 74–76
foci of, 71
hyperbolic, 72, 81–90
parabolic, 72, 77–81
standard forms for, 72–73, 74–75, 77–79, 82–83, 86–87
- Constant. *See also* Constant of integration
absolute, 6
arbitrary, 6
numerical, 6
parametric, 6, 33
- Constant of integration, 388–389. *See also* Boundary conditions; Initial conditions; Integration
- Constant product curve, 349–351. *See also* Equal product curve; Isoquant; Production function
- Constrained optima, 367–383, 607–613. *See also* Kuhn-Tucker conditions; Lagrange multipliers

- Constraints. *See also*
 Constrained optima;
 Linear programming
 in the dual problem, 631-632
 in the primal problem,
 620-621, 624-625, 629
- Consumers' surplus, 418-422.
See also Producers'
 surplus
- Consumption function, 46-48,
 172-174, 506-507. *See also*
 Consumption model
 marginal propensity to
 consume, 172, 396-397
 marginal propensity to save,
 172, 396-397
 multiplier, 172
- Consumption model, 506-507,
 525-526. *See also*
 Consumption function
- Continuity, 151-159, 324. *See also*
 Discontinuity;
 Smooth function
 and differentiability, 217-219
 in an interval, 151
 at a point, 151
- Convergence of an infinite series,
 298-299. *See also* Infinite
 series; Interval of
 convergence; Maclaurin's
 series; Power series;
 Radius of convergence;
 Taylor's series
 absolute, 299-300
 conditional, 299-300
 tests for
 alternating series, 299-300
 Cauchy ratio, 299-300
 comparison series, 299-300
 necessary condition, 299-
 300
- Convergence of an integral, 402
- Convex function, 611
- Coordinates of a point, 16. *See also*
 Polar coordinates;
 Rectangular coordinates
 abscissa, 17
 ordinate, 17
 polar, 129-131
- Cosine, 125, 127, 130. *See also*
 Trigonometric function
- Cost, 44-46, 168-169, 245-254,
 340-341, 394-395, 425-
 430, 605-606. *See also*
 Cost function
 average, 168, 245-246, 394-
 395
 fixed, 44-46
 joint, 340-341
 marginal, 168-169, 245-246,
 340-341, 394-395
 total, 44-46, 168, 245-246,
 394-395, 605
 variable, 44-46
- Cost function, 44-46, 168-169,
 245-254, 425-430,
 605-606. *See also* Cost
 cubic, 248-250
 exponential, 252-254
 linear, 246-247
 polynomial, 251-252
 quadratic, 247-248
- Cotangent, 125. *See also*
 Trigonometric function
- Cramer's Rule, 589-590. *See also*
 Simultaneous linear
 equations, solution of
- Curve. *See also* Family, of
 curves; Function
 algebraic, 50
 asymptotes of, 57-59
 exponential, 111-114
 higher plane, 50
 intercepts of, 51
 logarithmic, 114-117
 nonlinear, 50-133
 polynomial, 50
 transcendental, 50
- Decreasing function, 219-220,
 230. *See also* First
 derivative; Function
- Definite integral, 399-417. *See also*
 Integration
 geometric interpretation of,
 399-401
- Degenerate conic, 71, 73, 75, 79,
 83. *See also* Conic
 section; Quadratic
 equation
- Degree. *See also* Difference
 equation; Differential
 equation; Equation
 of an equation
 algebraic, 18, 50
 difference, 511
 differential, 470
 of a polynomial, 50
 of a term, 18, 50
 of a variable, 18
- Degree measure, 124-128. *See also*
 Trigonometric
 function
- Demand function, 37-40, 43-44,
 92-98, 170, 257-265,
 341-344. *See also*
 Demand surface; Market
 equilibrium; Supply
 function
 linear, 37-40, 43-44
 quadratic, 92-98
- Demand surface, 341-344. *See also*
 Competitive
 commodities;
 Complementary
 commodities; Partial
 demand; Partial
 differentiation
- Dependent variable, 8-9, 141,
 324. *See also* Function;
 Variable
- Derivative. *See also*
 Differentiable function;
 Differentiation;
 Maximum; Minimum;
 Point of inflection; Rate
 of change of a function
 first, 161-168, 219-230, 240-
 244
 geometric interpretation of,
 161-162, 219-221
 higher-order, 207-212
 partial, 324-328
 rules for obtaining, 174-202,
 317-321
 algebraic function, 174-177,
 201, 317-318
 composite function, 196-
 197, 202, 321, 334-
 335
 exponential function, 187-
 190, 202, 319
 implicit function, 213-215,
 335-338
 inverse function, 194-195,
 202, 320-321
 logarithmic function, 185-
 187, 202, 318-319
 polynomial function, 177,
 201, 317
 trigonometric function, 191-
 193, 202, 319-320
 second, 231-244
 total, 332-335
- Descartes, René, 16
- Determinant of a matrix, 564-
 569. *See also* Cramer's
 Rule; Inverse of a
 matrix; Simultaneous
 linear equations
 rules for obtaining
 diagrammatic, 565
 expansion by cofactors,
 565-566
- Diagonal matrix, 551-552. *See also*
 Identity matrix
- Difference equation, 509-540.
See also Cobweb model;
 Consumption model;
 Harrod model; Metzler
 inventory model;
 Samuelson's interaction
 model
 classification of
 degree, 510-511
 order, 510-511
 general solution of, 512-514
 linear, 511-512
 particular solution of, 512-514
 solution of linear first order,
 515-518
 solution sequence, behavior
 of, 518-521;
 equilibrium of, 522;
 stability of, 522;
 stationary value of,
 522
 solution of linear second
 order, 528-531
 nonhomogeneous, 533-535
 solution sequence, behavior
 of, 531-532;
 equilibrium of, 535-
 536; stability of,
 535-536; stationary
 value of, 535-536

- Differentiability and continuity, 217-219
- Differentiable function, 162. *See also* Differential; Differentiation
- Differential, 203-206. *See also* Derivative; Differentiation
- partial, 330-331
- total, 330-331
- of x , 203-204
- of y , 203-204
- Differential equation, 470-507. *See also* Consumption model; Domar debt models; Domar macro model; Evans price adjustment model
- classification of, 470-471
- degree, 470
- order, 470
- of first order and first degree, 476-498
- exact, 477, 487-491
- homogeneous, 477, 480-484
- linear, 477, 491-495
- linear in a function of x or y , 477, 495-498; Bernoulli equations, 498
- separable, 477-480
- general solution of, 471-475
- of higher order and/or degree, 498-499
- ordinary, 470
- partial, 470
- particular solution of, 471-475
- Differentiation, rules for, 174-202, 332-338. *See also* Derivative; Differentiable function; Implicit differentiation
- algebraic function, 174-177, 201, 317-318
- composite function, 196-197, 202, 321, 334-335
- exponential function, 187-190, 202, 319
- implicit function, 213-215, 335-338
- inverse function, 194-195, 202, 320-321
- logarithmic function, 185-187, 202, 318-319
- polynomial function, 177, 201, 317
- summary of, 201-202
- summary of procedure of, 199-200
- trigonometric function, 191-193, 202, 319-320
- Directrix of a conic, 71. *See also* Conic section
- Discontinuity, 151-159. *See also* Continuity; Discontinuous function; Discrete function
- finite, 152
- infinite, 152
- missing point, 152
- Discontinuous function, 151-159. *See also* Discontinuity; Discrete function
- Discrete function, 155-159. *See also* Discontinuity; Discontinuous function
- Distribution, marginal
- productivity theory of, 346-349
- Distribution of income, Pareto's law, 107-109
- Disturbance term, 363, 598. *See also* Regression analysis
- Domain. *See also* Function; Relation
- of a function, 8
- of a relation, 8
- Domar debt models, 503-505
- Domar macro model, 502-503
- Dominance, 651-653. *See also* Game theory; Matrix game; Minimax solution; Saddle point
- Double integral, 459-460. *See also* Integration; Multiple integral; Probability
- Dual, 631-635. *See also* Linear programming; Primal
- Eccentricity of a conic, 71. *See also* Conic section
- Economic model, 502-507, 537-540. *See also* Cobweb model; Consumption model; Domar debt models; Domar macro model; Evans price adjustment model; Harrod model; Metzler inventory model; Samuelson's interaction model; Simultaneous linear equations
- dynamic, 502
- static, 502
- Elasticity, 171, 256-265, 266-271, 344-345. *See also* Logarithmic derivative; Marginal analysis, of revenue; Partial differentiation
- arc, 256, 258-259
- constant, 263-265
- cross, 258-263
- of demand, 266-271, 344-350
- partial, 344-350
- point, 171, 256-258
- Elementary operations. *See also* Determinant of a matrix; Inverse of a matrix
- on columns, 574
- on rows, 574
- Ellipse, 72, 74-76
- center of, 74
- major axis of, 74-75
- minor axis of, 74-75
- standard form of, 74-75
- Endogenous variable, 502. *See also* Economic model; Exogenous variable
- Equal product curve, 349-351. *See also* Constant product curve; Isoquant
- Equation. *See also* Curve; Difference equation; Differential equation; Function; Simultaneous linear equations
- graphical representation of, 16-133
- locus of, 18
- Equations. *See also* Economic model; Simultaneous linear equations
- consistent, 27-28
- independent, 27-28
- reduced form, 502
- simultaneous linear, 27-32, 589-592
- structural, 502
- Equilibrium. *See also* Demand function; Difference equation; Market equilibrium; Stationary value; Supply function
- of a difference equation, 522, 535-536
- in purely competitive markets, 42-44, 93-98
- Error, 204-206. *See also* Approximation; Differential equation
- percentage, 206
- relative, 206
- Euler's theorem, 348-349. *See also* Marginal productivity theory of distribution; Partial differentiation
- Evans price adjustment model, 505-506
- Exact differential equation, 477, 487-491
- Exogenous variable, 502. *See also* Economic model; Endogenous variable
- Expansion path, 351. *See also* Returns to outlay; Returns to scale; Ridge line; Scale line
- Explicit function, 351. *See also* Function
- Exponential cost function, 252-254. *See also* Cost function
- Exponential function, 111-114, 117-122, 187-190, 202, 252-254, 319. *See also* Function; Power function; Scientific notation

- Extent, 56–57. *See also*
Graphing, general
methods of
- Factorization, 60–61, 62. *See also*
Graphing, general
methods of
- Fair game, 641. *See also* Matrix
game; Value of a game
- Family. *See also* Parameter
of curves, 90
of straight lines, 33–36
- Finite discontinuity, 152. *See
also* Discontinuity
- Finite game, 640. *See also* Game
theory; Matrix game
- First derivative, 161–230. *See
also* Derivative;
Maximum; Minimum
interpretation of, 165–168
- First difference, 510. *See also*
Difference equation
- Foci of a conic, 71. *See also*
Conic section
- Frequency function, 414–417,
464–467. *See also* Area;
Joint frequency function;
Probability
- Function, 8–10, 323–324. *See
also* Consumption
function; Cost function;
Curve; Demand function;
Growth function;
Joint-cost function;
Production function;
Rate of change of a
function; Relation;
Revenue function;
Supply function;
Variable
algebraic, 50, 174–177, 201,
317–318
composite, 10, 196–197, 202,
321, 334–335
concave, 611
continuous, 151, 159
convex, 611
decreasing, 219–220, 230
dependent variable of, 8–9, 141,
324
differentiable, 162
discontinuous, 151–159
discrete, 155–159
domain of, 8
explicit, 213
exponential, 111–114, 117–122,
187–190, 202, 252–254,
319
frequency, 414–417, 464–467
Gompertz, 120–121
growth, 119–122
homogeneous, 348–349, 477,
480–484
implicit, 213–215, 335–338
increasing, 219–220, 230
independent variable of, 8–9,
141, 324
integrable, 400
inverse, 13–14, 55, 117, 194–
195, 202, 320–321
linear, 17–25
logarithmic, 114–117, 185–
187, 202, 318–319
nonlinear, 50–133
polynomial, 50, 177, 201, 317
power, 108
quadratic, 70–90
range of, 8
smooth, 218–219
step, 155–159
transcendental, 50
trigonometric, 123–131, 133,
191–193, 202
value of, 9
- Function value, 9. *See also*
Function
- Game matrix, 640–641. *See also*
Dominance; Game
theory; Matrix game;
Minimax solution; Payoff
matrix; Saddle point;
Strategies; Value of a
game
- Game theory, 639–657. *See also*
Dominance; Game
matrix; Matrix game;
Minimax solution; Payoff
matrix; Saddle point;
Strategies; Value of a
game
- Gaussian elimination, 574–577.
See also Inverse of a
matrix
- General solution. *See also*
Constant of integration;
Particular solution
of a difference equation,
512–514
of a differential equation,
471–475
- Generic constant, 472. *See also*
Constant of integration;
Differential equation
- Gompertz function, 120–121.
See also Function;
Growth function
- Graph, 18. *See also* Equation;
Locus
- Graphical solution of a matrix
game, 653–656. *See also*
Game theory; Matrix
game; Minimax solution
- Graphing. *See also* Asymptote;
Extent; Factorization;
Intercepts; Maximum;
Minimum; Point of
inflection; Symmetry
general methods of, 50–70
using first and second
derivatives, 219–244
- Growth function, 119–122. *See
also* Curve; Function
biological, 119–120
Gompertz, 120–121
learning, 121–122
- Harrod model, 523–524
- Hessian determinant, 605, 607.
See also Bordered
Hessian determinant;
Maximum of a function
of n variables
- Higher-order derivative, 207–
212. *See also* Derivative
- Homogeneous differential
equation, 477, 480–484.
See also Differential
equation
- Homogeneous function, 348–349,
477, 480–484
- Hyperbola, 72, 81–90. *See also*
Conic section
asymptotes of, 82–83, 86
center of, 81
equilateral, 82, 86–90
generalized equilateral, 87–90
of Fermat, 87
rectangular, 82, 86–90
standard form of, 82–83, 86
transverse axis of, 81
- Identity, trigonometric, 129. *See
also* Trigonometric
function
- Identity matrix, 552–553. *See
also* Diagonal matrix;
Inverse of a matrix
- Imaginary locus, 61–62. *See also*
Locus
- Implicit differentiation, 213–215.
See also Differentiation
- Implicit function, 213–215, 335–
338. *See also* Function;
Implicit differentiation
- Implicit prices, 635. *See also*
Dual; Imputed prices;
Shadow prices
- Improper integral, 402. *See also*
Convergence;
Divergence; Integral
- Imputed prices, 635. *See also*
Dual; Implicit prices;
Shadow prices
- Income, 46–48, 172–174, 396–
397, 506–507, 525, 526.
See also Consumption
function; Consumption
model
- Increasing function, 219–220,
230. *See also* First
derivative; Function
- Indefinite integral, 388–398.
See also Constant of
integration; Integration
interpretation of, 388–389
- Independent variable, 8–9, 141,
324. *See also* Function;
Variable
- Indeterminate expression, 7, 20.
See also Indeterminate
form; Undefined
expression
- Indeterminate form, 148–149,
289–296. *See also*
Discontinuity;

- Indeterminate form (*Continued*)
 Indeterminate expression;
 L'Hospital's rule; Limit
 Type $\frac{0}{0}$, 148–149, 289–290, 296
 Type $\frac{\infty}{\infty}$, 148–149, 289–290, 296
 Type $\infty \cdot 0$, 291–295, 296
 Type 1^∞ , 291–295, 296
 Type 0^0 , 291–295, 296
 Type ∞^0 , 291–295, 296
 Type $\infty - \infty$, 295, 296
- Indifference curve, 353. *See also*
 Indifference map; Utility
- Indifference map, 353. *See also*
 Indifference curve;
 Utility
- Infinite discontinuity, 152. *See also*
 Discontinuity
- Infinite games, 640. *See also*
 Game theory; Matrix
 game
- Infinite sequence, 298. *See also*
 Difference equation,
 solution sequence of;
 Infinite series;
 Recursion formula
- Infinite series, 298–315. *See also*
 Convergence of an infinite
 series; Infinite sequence;
 Maclaurin's series;
 Recursion formula;
 Taylor's series
 alternating, 299–300
 convergent, 298–299
 divergent, 298–299
 oscillating, 298, 520
- Inflection, point of, 236–239, 240.
See also Second
 derivative
- Initial conditions. *See also*
 Boundary conditions;
 Constant of integration;
 Particular solution
 for a difference equation, 512
 for a differential equation,
 472
 for integration, 389
- Inner product of two vectors,
 548. *See also* Vector
- Input-output analysis, 614–618
- Integrable function, 400. *See also*
 Integration
- Integral, 388. *See also*
 Integration
 definite, 402
 double, 459–461, 464–467
 improper, 402
 indefinite, 388
 multiple, 459–463
 particular, 388
- Integrand, 388. *See also*
 Integration
- Integrating factor, 491–492. *See also*
 Differential equation
- Integration, 387–467. *See also*
 Constant of integration;
 Integral; Multiple
 integration
 approximate methods of, 449–
 458
- Simpson's rule, 449, 451–
 453, 455–458
- Taylor's expansion, 449,
 453–458
- trapezoidal rule, 449–451,
 455–458
- definite, 399–429
 double, 459–461, 464–467
 fundamental theorem of,
 404
 indefinite, 388–392
 multiple, 459–463
 partial, 459
 by partial fractions, 438–442
 by parts, 435–437
 by rationalization, 443–447
 by standard forms, 390–392,
 431–434
 by substitutions, 443–448
- Intercepts
 of a curve, 51
 of a straight line, 18–19
- Interest. *See also* Compound
 interest; Simple interest
 compound, 118, 521
 simple, 521
- Interval of convergence, 304. *See also*
 Infinite series; Power
 series; Radius of
 convergence
- Inventory models, 285–288,
 539–540
- Inverse of a function, 13–14,
 55, 117, 194–195, 202,
 320–321. *See also*
 Function
- Inverse of a matrix, 569–584. *See also*
 Determinant of a
 matrix; Elementary
 operations; Game
 matrix; Gaussian
 elimination; Identity
 matrix; Input-output
 analysis; Linear
 programming; Markov
 chain process;
 Simultaneous linear
 equations, solutions of
 by adjoint and determinant,
 577–579
 by canonical procedure using
 row or column
 operations, 574–577
 of a partitioned matrix,
 580–582
- Inverse of a relation, 13–14. *See also*
 Relation
- Isocost curve, 351
- Isoquant, 349–351. *See also*
 Constant product curve;
 Equal product curve
- Joint-cost function, 340–341. *See also*
 Cost function
- Joint frequency function, 414–
 417, 464–467. *See also*
 Area; Frequency
 function; Probability
- Kuhn-Tucker conditions, 375–
 380, 611–613. *See also*
 Constrained optima
- Lagrange multipliers, 368–375,
 607–611. *See also*
 Constrained optima
- Law of diminishing marginal
 productivity, 346
- Least squares estimates, 363–364,
 598–599. *See also*
 Regression analysis
- Leibnitz, Gottfried, 139
- Leontief matrix, 615. *See also*
 Input-output analysis
- L'Hospital's rule, 289–296. *See also*
 Indeterminate form;
 Limit
- Limit, 141–149. *See also*
 Asymptote; Convergence;
 Divergence;
 Indeterminate form;
 L'Hospital's rule
 of a function, 141–149
 left-hand, 144–146
 properties of, 146–149
 right-hand, 144–146
 of a sequence, 298
 of a series, 298
 of a variable, 141–142
- Limits of integration, 402. *See also*
 Definite integral
- Linear dependence, 585–588.
See also Rank;
 Simultaneous linear
 equations
- Linear difference equation, 511–
 540. *See also* Difference
 equation
- Linear differential equation, 477,
 491–498. *See also*
 Differential equation
- Linear equation, 17–25. *See also*
 Linear function;
 Simultaneous linear
 equations; Straight line
- Linear function, 17–25. *See also*
 Function; Linear
 equation; Maximization;
 Minimization; Straight
 line; Vector
 differentiation
- Linear programming, 619–635.
See also Constrained
 optima; Inverse of a
 matrix; Simplex method
 artificial variables of, 625–626,
 629, 633
 constraints of, 620–621, 624–
 625, 629, 631–632
 dual of, 631–635
 feasible solution of, 621, 624–
 625
 geometric solution of, 621–625
 of a matrix game, 656–657
 objective function of, 620–621,
 631–632
 primal of, 631–632

- Linear programming
(Continued)
primary variables of, 620
simplex solution of, 625–631
slack variables of, 625–626, 629, 633
- Locus, 18. *See also* Equation; Graph; Imaginary locus; Point locus
- Logarithm, 114–115. *See also* Base e ; Logarithmic function
Briggsian, 114
common, 114
Napierian, 114
natural, 114
- Logarithmic derivative, 257.
See also Elasticity; Logarithm
- Logarithmic function, 114–117, 117–122, 185–187, 202.
See also Function; Logarithm
- Maclaurin's series, 308. *See also* Infinite series; Taylor's series
- Majorize, 651. *See also* Dominance
- Marginal analysis. *See also* First derivative
of consumption, 172, 396–397
of cost, 168–169, 245–246, 340–341, 394–395
of demand, 341–344
of production, 346–349, 605
of revenue, 170, 266–271, 395–396, 425–429
of utility, 352–354, 380–383
- Marginal product, 346–349, 602.
See also Euler's theorem; Marginal productivity theory of distribution; Production function
- Marginal productivity theory of distribution, 349. *See also* Euler's theorem; Marginal product
- Marginal propensity to consume, 172–174, 396–397. *See also* Consumption function; Consumption model
- Marginal propensity to save, 172–174, 396–397. *See also* Consumption function; Consumption model
- Marginal rate of substitution, 354. *See also* Equal product curve
- Marginal utility, 352–354, 380–383. *See also* Utility function
- Market equilibrium, 42–44, 93–98. *See also* Demand function; Supply function
- Markov chain process, 661–666.
See also Inverse of a matrix; Matrix equilibrium of, 662–666
nonstationary, 662
states of, 661
stationary, 662
steady state of, 662–666
transition matrix of, 661
transition probabilities of, 661
- Mathematical programming, 620. *See also* Linear programming
- Matrix, 541–542. *See also* Determinant of a matrix; Elementary operations; Inverse of a matrix; Matrix operations
adjoint of, 566, 577, 647–650
cofactors of, 566
determinant of, 564–569
diagonal, 551–552
elements of, 541
idempotent, 555
identity, 552–553
inverse of, 569–584
left, 570
right, 570
minors of, 566
nonsingular, 570
null, 553–554
order of, 541–542
partitioned, 558–562, 580–582
principal minors of, 603
rank of, 585–588
singular, 570
square, 541
symmetric, 555
transpose of, 554–557, 558–559
- Matrix game, 639–657. *See also* Game matrix; Game theory; Minimax solution; Mixed strategy; Saddle point
dominance in, 651–653
graphical solution of, 653–656
linear programming solution of, 656–657
saddle point of, 641–642
solution of 2×2 games, 642–650
solution of $2 \times n$ and $m \times 2$ games, 650–656
solution of larger games, 656–657
- Matrix operations, 544–551, 558–562. *See also* Elementary operations; Inner product of two vectors; Matrix addition, 544–545, 559–560
multiplication, 546–551, 560–562
multiplication by a scalar, 545–546
subtraction, 544–545, 559–560
- Maximin solution, 639. *See also* Game theory; Matrix game; Minimax solution
- Maximum. *See also* Kuhn-Tucker conditions; Lagrange multipliers
of a function of n variables
constrained, 607–613
unconstrained, 602–606
of a function of one variable
absolute, 228–230
relative, 220–228, 230, 235
of a function of two variables
constrained, 367–383
unconstrained, 356–366
- Metzler inventory model, 539–540
- Minimax solution, 639. *See also* Dominance; Game theory; Matrix game; Maximin solution; Mixed strategy; Saddle point
- Minimum
of a function of n variables
constrained, 607–613
unconstrained, 602–606
of a function of one variable
absolute, 228–230
relative, 220–228, 230, 235
of a function of two variables
constrained, 367–383
unconstrained, 356–366
- Minorize, 651. *See also* Dominance
- Minors of a matrix, 566. *See also* Matrix; Principal minors of a matrix
- Missing-point discontinuity, 152.
See also Discontinuity
- Mixed partial derivative, 327.
See also Partial derivative; Second derivative
- Mixed strategy, 644. *See also* Dominance; Game theory; Matrix game; Minimax solution; Saddle point
- Monopoly
profit under, 275–279
taxation under, 280–284
- Morganstern, Oskar, 639
- Multiple integral, 459–463. *See also* Integral; Multiple integration
- Multiple integration, 459–463.
See also Integration; Multiple integral
- Multiplier, 172. *See also* Consumption function; Consumption model
- Multivariate regression analysis, 598–600. *See also* Least squares estimates; Regression analysis
- Napier, John, 114
- Net investment, 397–398. *See also* Capital formation; Integration
- Newton, Sir Isaac, 139

- Nonlinear curve, 50–133. *See also* Curve; Function
- Nonsingular matrix, 570. *See also* Determinant of a matrix; Inverse of a matrix; Singular matrix
- Nonstrictly determined game, 643–645. *See also* Game matrix; Matrix game; Minimax solution; Mixed strategy
- Non-zero-sum game, 640. *See also* Game theory; Matrix game; Zero-sum game
- Null matrix, 553–554. *See also* Matrix
- Numbers
complex, 7
imaginary, 7
real, 7
- Objective function, 620, 631–632. *See also* Linear programming
- Operator, 510
- Ordinate, 17. *See also* Rectangular coordinates
- Oscillating series, 298, 520. *See also* Difference equation, solution sequence of; Infinite series
- Parabola, 72, 77–81. *See also* Conic section
axis of, 77
orientation of, 77–79
standard form of, 77–79
vertex of, 77
- Parameter, 6, 33. *See also* Constant, parametric; Family, of curves
- Pareto, Vilfredo, 107
- Pareto's law, 107–109. *See also* Power function; Scientific notation
- Partial demand, 341–344. *See also* Demand surface
- Partial derivative, 324–328. *See also* Derivative; Mixed partial derivative; Partial differentiation
- Partial differential, 330–331. *See also* Differential
- Partial differentiation, 324–328. *See also* Differentiation; Partial derivative
- Partial elasticities of demand, 344–345. *See also* Demand surface; Elasticity; Partial differentiation
- Particular integral, 388. *See also* Boundary conditions; Constant of integration; Initial conditions
- Particular solution. *See also* Boundary conditions; Constant of integration; Initial conditions
of a difference equation, 512–514
of a differential equation, 471–475
- Payoff matrix, 640–641. *See also* Game matrix; Matrix game
- Period analysis, 509. *See also* Difference equation
- Point locus, 61–62. *See also* Locus
- Point of inflection, 236–244. *See also* Second derivative
- Polar coordinates, 129–131. *See also* Coordinates of a point; Function; Trigonometric function
- Polynomial, 50. *See also* Polynomial function
- Polynomial function, 50, 177, 201, 317. *See also* Function, polynomial
- Postmultiply, 547. *See also* Matrix operations, multiplication; Premultiply
- Power function, 111. *See also* Exponential function; Function
- Power series, 304–315. *See also* Infinite series; Maclaurin's series; Taylor's series
- Premultiply, 547. *See also* Matrix operations, multiplication; Postmultiply
- Primal, 631–632. *See also* Dual; Linear programming
- Principal minors of a matrix, 603. *See also* Matrix; Minors of a matrix
- Probability, 414–417, 464–467. *See also* Area; Double integral; Frequency function; Integral
- Producers' surplus, 418, 422–425. *See also* Consumers' surplus
- Production function, 346–349, 351–352, 605. *See also* Euler's theorem; Function; Marginal product
linear homogeneous, 348–349
- Product transformation curve, 98–104
- Properties of limits, 146–149. *See also* Indeterminate form; L'Hospital's rule; Limit
- Quadratic form, 595–597. *See also* Vector differentiation
- Quadratic function, 70–90. *See also* Conic section; Function
general equation of, 70–71
- Radian measure, 124–128. *See also* Trigonometric function
- Radius of convergence, 304. *See also* Infinite series; Interval of convergence; Power series
- Range. *See also* Function; Relation; Variable
of a function, 8
of a relation, 8
of a variable, 6
- Rank of a matrix, 585–588. *See also* Linear dependence; Matrix; Simultaneous linear equations
- Rate of change of a function. *See also* Derivative
average, 167–168
instantaneous, 167–168
- Rate of technical substitution, 349. *See also* Equal product curve
- Reciprocal substitution, 447–448. *See also* Integration, by substitution
- Rectangular coordinates, 16–17. *See also* Coordinates of a point; Polar coordinates
axes of, 16–17
origin of, 16
plotting points in, 17
quadrants of, 16–17
- Recursion formula, 298. *See also* Infinite sequence; Infinite series
- Reduced form equations, 502. *See also* Economic model; Structural equations
- Regression analysis, 363–364, 598–599. *See also* Bivariate regression analysis; Least squares estimates; Multivariate regression analysis
- Relation, 8. *See also* Function; Variable
binary, 8
dependent variable of, 8
domain of, 8
independent variable of, 8
inverse, 8
range of, 8
- Relative maximum, 220–228, 230, 235. *See also* Maximum
- Relative minimum, 220–228, 230, 235. *See also* Minimum
- Returns to outlay, 352. *See also* Equal product curve; Expansion path; Returns to scale; Ridge line; Scale line
- Returns to scale, 351–352. *See also* Equal product curve; Expansion path; Returns

- Returns to scale (*Continued*)
to outlay; Ridge line;
Scale line
- Revenue, from taxation, 271–275.
See also Revenue function
- Revenue, 44–46, 170–171, 266–271, 395–396, 425–429, 605. *See also* Demand; Elasticity
average, 170, 266–267, 395–396
and elasticity of demand, 266–271
marginal, 170, 266–271, 395–396
total, 170, 266, 395–396, 605
- Ridge line, 350. *See also* Equal product curve; Expansion path; Returns to outlay; Returns to scale; Scale line
- Saddle point, 357–358, 641–642.
See also Game matrix; Matrix game; Maximum; Minimax solution; Minimum; Second derivative
- Samuelson's interaction model, 537–538
- Scalar, 545. *See also* Matrix
- Scale line, 351–352. *See also* Equal product curve; Expansion path; Returns to outlay; Returns to scale; Ridge line
- Scientific notation, 112
- Secant, 125, 161–162. *See also* Trigonometric function
- Second derivative, 207, 231–244.
See also Concavity; Derivative; Partial derivative; Point of inflection
interpretation of, 231–232
- Separable differential equation, 477–480. *See also* Differential equation
- Sequence, 298. *See also* Infinite sequence
- Series, 298–315. *See also* Infinite series
- Set, 2–5. *See also* Venn Diagram
complement, 4
defining relation of, 2
difference, 4
element of, 2
empty, 3
intersection, 3
proper subset, 3
subset, 3
union, 3
universal, 4
universe, 4
- Shadow prices, 635. *See also* Dual; Implicit prices; Imputed prices
- Simple interest, 521. *See also* Compound interest; Interest
- Simplex method, 625–631. *See also* Inverse of a matrix; Linear programming
steps in solution by, 626
tableau of, 625–626
variations in, 631
- Simpson's rule, 449, 451–453, 455–458. *See also* Integration, approximate methods of
- Simultaneous linear equations, 26–32, 585–592. *See also* Economic model; Linear dependence; Rank
consistent, 27, 28, 587–588
dependent, 27, 28, 585–588
inconsistent, 27, 28, 587–588
independent, 27, 28, 585–588
solution of, 28–32, 589–592
by elimination, 29–32
by matrix algebra, 29–32, 589–592; Cramer's rule 589–590; inversion procedures, 589
by substitution 29–32
- Sine, 125, 127, 130. *See also* Trigonometric function
- Singular matrix, 570. *See also* Determinant of a matrix; Inverse of a matrix; Nonsingular matrix
- Slack variables, 625–626, 629, 633. *See also* Linear programming; Simplex method
- Slope of a line, 19–20, 161. *See also* First derivative; Rate of change of a function; Tangent to a curve
- Smooth function, 218–219. *See also* Continuity; Function; Second derivative
- Solution. *See also* Difference equation; Differential equation; Equation; Simultaneous linear equations
of a difference equation
general, 512–514
particular, 512–514
of a differential equation
general, 471–475
particular, 471–475
of simultaneous linear equations, 28–32, 589–592
- Square matrix, 541. *See also* Matrix
- Standard forms for conics, 72–73, 74–75, 77–79, 82–83, 86–87. *See also* Circle; Conic section; Ellipse; Hyperbola; Parabola
- Standard forms for integration, 390–392, 431–434. *See also* Integration
- States of a Markov chain
process, 661. *See also* Markov chain process
- Stationary probability distribution, 662. *See also* Markov chain process
- Stationary value, 535–536. *See also* Difference equation; Equilibrium
of a difference equation, 535–536
- Steady state, 662–666. *See also* Equilibrium; Markov chain process
- Step function, 155–159. *See also* Function
- Straight line. *See also* Linear equation; Straight lines
general equation of, 18, 25
horizontal, 25
intercept form of, 22–23, 25
intercepts of, 18–19
point-slope form of, 23–24, 25
slope-intercept form of, 24, 25
two-point form of, 21–22, 25
vertical, 24, 25
- Straight lines, 17–36. *See also* Simultaneous linear equations; Straight line
families of, 33–35
intersecting, 27–32
parallel, 26
perpendicular, 26
- Strategies, 640. *See also* Game matrix; Matrix game; Minimax solution; Saddle point
mixed, 641–642
pure, 644–646
- Strictly determined game, 641.
See also Game matrix; Matrix game; Minimax solution; Saddle point
- Structural equations, 502. *See also* Economic model; Reduced form equations
- Supply function, 40–42, 43–44, 92–98, 170, 257–265. *See also* Demand function; Market equilibrium
linear, 40–42, 43–44
quadratic, 92–98
- Symmetric matrix, 555. *See also* Matrix
- Symmetry
with respect to an axis, 51–55
with respect to the origin, 51–55
with respect to a point, 55–56
with respect to a straight line, 55–56
with respect to the x-axis, 51–55
with respect to the y-axis, 51–55
- Tangent of an angle, 19–21, 125, 127, 130, 161, 204, 326.
See also Angle of inclination; First

- Tangent of an angle (*Continued*)
 derivative: Slope of a line; Trigonometric function
- Tangent to a curve, 162. *See also* First derivative, Rate of change of a function; Slope of a line
- Taxation
 maximization of monopolistic profit under, 280-284
 maximization of revenue from, 271-275
- Taylor's series, 308-315. *See also* Infinite series; Maclaurin's series
- Taylor's theorem, 308-315, 321-322, *See also* Taylor's series
- Technological matrix, 615. *See also* Input-output analysis; Leontief matrix
- Total derivative, 332-335. *See also* Derivative
- Total differential, 330-331. *See also* Differential
- Transcendental function, 50. *See also* Function
- Transition probabilities, 661. *See also* Markov chain process
- Transpose of a matrix, 554-557, 558-559. *See also* Matrix
- Trapezoidal rule, 449-451, 455-458. *See also* Integration, approximate methods of
- Trigonometric function, 123-131, 133, 191-193, 202. *See also* Function
 cosecant, 125-129
 cosine, 125-129
 identities, 129
 polar coordinates, 129-131
 radian measure of, 124-125
 secant, 125-129
 sine, 125-129
 tangent, 125-129, 326
- Undefined expression, 7, 20. *See also* Indeterminate expression; Indeterminate form
- Undetermined multipliers, 368-375, 607-611. *See also* Lagrange multipliers
- Utility function, 352-354, 610-611. *See also* Marginal analysis, of utility
 constrained maximization of, 380-383
- Value of a game, 641, 645. *See also* Game matrix; Matrix game, Minimax solution; Strategies
- Variable, 6-7, 8. *See also* Function; Relation
 continuous, 6
 dependent, 8-9, 141, 324
 discrete, 6
 independent, 8-9, 141, 324
 range of, 6
- Vector, 542-543. *See also* Inner product of two vectors; Matrix
 column, 542
 components of, 542-543
 dimension of, 542-543
 of functions, 595
 row, 543
- Vector differentiation, 593-600. *See also* Maximization; Minimization; Partial differentiation; Vector
 of a bilinear form, 597-598
 of a linear function, 593-595
 of a quadratic form, 595-597
 of a vector of functions, 595
- Venn Diagram, 4. *See also* Set
- Von Neumann, John, 639
- Zero-sum game, 640. *See also* Game theory; Matrix game; Non-zero-sum game

